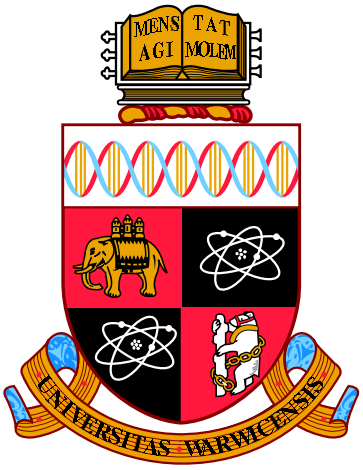

Measurement of the $B^0_{d,s} \to K^{*\pm}h^{\mp}$ decay branching fractions at the LHCb experiment.

by

David Dossett

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Physics

October 2013

THE UNIVERSITY OF WARWICK

Contents

List of Tables

List of Figures

Acknowledgments

Firstly I would like to thank Tim Gershon, for his boundless generosity in giving me this opportunity, as well as his reserves of patience while this document was being written. Tom Latham also deserves enormous praise for his input and guidance during this analysis. Both Tim and Tom are incredible sources of expertise, without which I could not have carried this research to conclusion.

The rest of the Warwick LHCb group has always been ready with experience and knowledge when needed, for which I am forever indebted. A special thank you is reserved for Matt Williams, who has helped me through his limitless knowledge of computer programming and being a true friend during the course of these studies. I would like to thank the members of the B2KShh working group, both past and present, who have made this study possible through their previous work.

My thanks goes out to my family, who have supported me without fault during this endeavour. I am indebted to my friend Andrew, who has managed to keep me both fed and sane as I wrote this document.

Abstract

Measurements of the branching fractions of $B^0_{d,s} \to K^{*\pm}h^{\mp}$ ($h = \pi, K$) decays are performed, using pp collision data at the LHCb experiment. The data sample corresponds to an integrated luminosity of $1.0\,fb^{-1}$, collected at a centre-of-mass energy of 7 TeV. First observation of the $B^0_s \to K^{*\pm}K^{\mp}$ and evidence for the $B^0_s \to K^{*-}\pi^+$ decay modes are reported. The following branching fraction measurements and limits of the $B^0_{d,s} \to K^{*\pm}h^{\mp}$ ($h = \pi, K$) decay modes, relative to the well measured $B^0 \to K^{*+}\pi^-$ decay, are obtained

$$\frac{\mathcal{B}\left(B^0_s \to K^{*\pm}K^{\mp}\right)}{\mathcal{B}\left(B^0 \to K^{*+}\pi^-\right)} = 1.49 \pm 0.22\,(\text{stat.}) \pm 0.18\,(\text{syst.})$$

$$\frac{\mathcal{B}\left(B^0 \to K^{*\pm}K^{\mp}\right)}{\mathcal{B}\left(B^0 \to K^{*+}\pi^-\right)} < 0.06 \quad \text{at } 95\%\,\text{CL.}$$

$$\frac{\mathcal{B}\left(B^0_s \to K^{*-}\pi^+\right)}{\mathcal{B}\left(B^0 \to K^{*+}\pi^-\right)} = 0.39 \pm 0.13\,(\text{stat.}) \pm 0.05\,(\text{syst.})$$

Introduction

The LHCb experiment was designed to perform precision measurements of the properties of b hadrons. Experimental tests of the Standard Model predictions can reveal discrepancies, which may then indicate the presence of new physics. While the Standard Model continues to be a tremendous success, it cannot yet explain certain important observations. Specifically, the dominance of matter over antimatter in the Universe is not yet fully explained by the amount of *CP*-violation observed. Measurements of the *CP*-violating decays in the B system may reveal larger sources of *CP*-violation than predicted by the current model.

In this thesis, I present measurements of the branching fractions of $\mathrm{B}^0_{d,s} \to \mathrm{K}^{*\pm}\mathrm{h}^{\mp}$ ($\mathrm{h} = \pi, \mathrm{K}$) decays. These channels are potentially sensitive to the effects of new physics, and, particularly in the case of the B^0_s decays, have not previously been studied in detail.

Prior to this analysis, I was involved with the measurements of the $\mathrm{B}^0_{d,s} \to \mathrm{K}^0_{\mathrm{S}}\mathrm{h}^{\pm}\mathrm{h}'^{\mp}$ relative branching fractions at LHCb [1], giving the results:

$$
\begin{aligned}
\frac{\mathcal{B}\left(\mathrm{B}^0 \to \mathrm{K}^0_{\mathrm{S}}\mathrm{K}^{\pm}\pi^{\mp}\right)}{\mathcal{B}\left(\mathrm{B}^0 \to \mathrm{K}^0_{\mathrm{S}}\pi^{+}\pi^{-}\right)} &= 0.128 \pm 0.017\,(\text{stat.}) \pm 0.009\,(\text{syst.}),\\
\frac{\mathcal{B}\left(\mathrm{B}^0 \to \mathrm{K}^0_{\mathrm{S}}\mathrm{K}^{+}\mathrm{K}^{-}\right)}{\mathcal{B}\left(\mathrm{B}^0 \to \mathrm{K}^0_{\mathrm{S}}\pi^{+}\pi^{-}\right)} &= 0.385 \pm 0.031\,(\text{stat.}) \pm 0.023\,(\text{syst.}),\\
\frac{\mathcal{B}\left(\mathrm{B}^0_s \to \mathrm{K}^0_{\mathrm{S}}\pi^{+}\pi^{-}\right)}{\mathcal{B}\left(\mathrm{B}^0 \to \mathrm{K}^0_{\mathrm{S}}\pi^{+}\pi^{-}\right)} &= 0.29 \pm 0.06\,(\text{stat.}) \pm 0.03\,(\text{syst.}) \pm 0.02(f_s/f_d),\\
\frac{\mathcal{B}\left(\mathrm{B}^0_s \to \mathrm{K}^0_{\mathrm{S}}\mathrm{K}^{\pm}\pi^{\mp}\right)}{\mathcal{B}\left(\mathrm{B}^0 \to \mathrm{K}^0_{\mathrm{S}}\pi^{+}\pi^{-}\right)} &= 1.48 \pm 0.12\,(\text{stat.}) \pm 0.08\,(\text{syst.}) \pm 0.12\,(f_s/f_d),\\
\frac{\mathcal{B}\left(\mathrm{B}^0_s \to \mathrm{K}^0_{\mathrm{S}}\mathrm{K}^{+}\mathrm{K}^{-}\right)}{\mathcal{B}\left(\mathrm{B}^0 \to \mathrm{K}^0_{\mathrm{S}}\pi^{+}\pi^{-}\right)} &\in [0.004; 0.068]\,\text{at}\,90\%\,\text{CL}.
\end{aligned}
$$

These measurements were performed on data taken during 2011 at LHCb, with an integrated luminosity of $1.0\,\mathrm{fb}^{-1}$. This work was started prior to the completion of those results and serves to analyse the same data, looking specifically at the

resonant structure in the low $K^0_S\pi$ invariant mass region. The previous analysis measured the total branching fractions of all intermediate states proceeding to a final state. Here, the branching fractions of the decays which proceed through an intermediate $K^*(892)^{\pm}$ resonance are measured.

Chapter 1 of this document describes the basic theory underpinning the Standard Model and the current status of the predictions for the relevant branching fractions. Chapter 2 then details the requirements and final design of the LHC and the LHCb detector. During the earlier stages of my Ph.D. studies, I was involved with research on the VELO sub-detector at LHCb. For example, I helped perform the mapping of the VELO material using material interactions from beam-gas events, see Section 2.3. This was used to compare the material budget of the real VELO with a simulation and to perform alignment studies.

Chapters 3-8 describe the different parts of the relative branching fraction analysis, culminating in the final calculation of the relative branching fractions of the $B^0_{d,s} \to K^{*\pm}h^{\mp}$ decays. The implications of the results and potential for further studies are given in Chapter 9.

1 Theory

1.1 Introduction

In this chapter, the fundamental particles and interactions of the Standard Model will be introduced and a description of how the Standard Model theory accounts for their properties will be given. The LHCb experiment's main focus is the precise measurement of *CP* violation in the b-quark sector. This is intimately linked to quark mixing in the charged current weak interactions as described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix. Therefore the theoretical reasons for these effects are discussed in Sections 1.3-1.4.

The basics of the Standard Model as a gauge theory and the types of Lagrangian terms needed will be explained. Then there will be a brief overview of the electroweak theory to lead into how the vector bosons acquire mass. This will then lead into how the fermion mass terms are constructed and the consequences for quark mixing. The resulting CKM matrix and its links to *CP* violation in the Standard Model can then be discussed. There will then be a brief description of the history of *CP* violation and its importance, including an overview of the different ways in which *CP* violation can occur.

Finally, the reasons for interest in the charmless decays of B mesons studied in this thesis, previous measurements of similar decays, and theoretical predictions

for the branching fractions of the $B_s^0 \to K^{*\pm}K^{\mp}$ modes will be given.

1.2 The Standard Model

The fundamental particles

In the Standard Model there are two distinct types of particles, fermions and bosons. Fermions have half-integer spins and obey Fermi-Dirac statistics, while bosons have integer spins and obey Bose-Einstein statistics. There are 12 fermions in the Standard Model, six leptons and six quarks. These can be further divided into the up-type quarks (u, c, t), the down-type quarks (d, s, b), the charged leptons (e, μ, τ) and finally the neutrinos (ν_e, ν_μ, ν_τ). The fermions are usually arranged into three generations, where the particles in each generation can be considered higher mass copies of the previous generation but with identical quantum numbers. The fermions, their charges and their generations are summarised in Fig. 1.1. Not shown is the corresponding anti-particle of each fermion, which is an oppositely charged particle with otherwise identical fundamental properties, such as mass and spin.

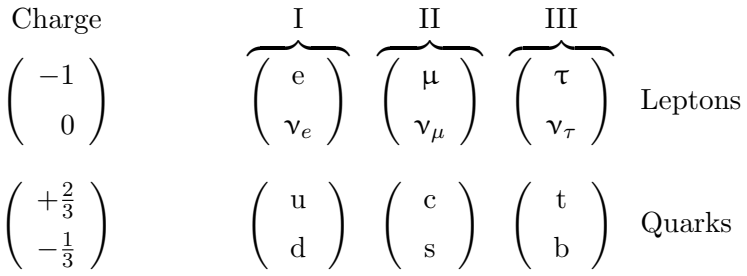

Figure 1.1: The three lepton and quark generations, split into their respective doublets with labelled charge values.

There are four types of gauge bosons in the Standard Model, which mediate the fundamental forces of nature, the photon (γ), the gluons (g), the $W^{\pm}$ and the Z^0. These bosons are exchanged between particles, allowing particles to interact with one another. However, not all of the bosons couple to all of the fermions. The photon only couples to charged particles and therefore cannot interact with the neutrinos, gluons, the Z^0 or even other photons. Gluons only couple to particles which carry the colour charge and so they only couple to the quarks and gluons themselves. The weak force can interact with all of the fermions and is therefore the only force in the Standard Model coupled to neutrinos.

There is a fifth boson, called the Higgs (H^0), which does not mediate a force but is instead linked to the appearance of mass in the Standard Model. Until

recently the Higgs boson had not been observed. Recently, the ATLAS and CMS experiments have observed a new boson at the predicted mass of the Higgs boson [2, 3]. The fundamental properties of the bosons are very different from one another, as summarised in Table 1.1.

Table 1.1: The bosons in the Standard Model and their respective spins, masses and interactions [2–4]. It should be noted that the spin of the observed Higgs boson could actually be either 0 or $2\hbar$, since the boson has an observed decay of $H^0 \to \gamma\gamma$. Recent measurements favour the expected 0-spin Higgs [5, 6].

Boson	Interaction	Mass (GeV/c^2)	Spin ($\hbar$)
photon (γ)	Electromagnetic	0	1
$W^{\pm}$	Weak Charged Current	80.385 ± 0.015	1
Z^0	Weak Neutral Current	91.1876 ± 0.0021	1
gluon (g)	Strong	0	1
Higgs (H^0)	Mass	$\sim$126	0

The fundamental forces

In the Standard Model the interactions of particles with the different forces are described by several different Quantum Field Theories (QFTs). The electromagnetic force governs the interactions of electrically charged particles. These are described by the theory of Quantum ElectroDynamics (QED). The strong force is responsible for the binding (and confinement) of quarks into hadrons, as well as the attraction between nucleons in the nucleus of an atom. Particles with colour charge can interact with the strong force and these interactions are described by Quantum ChromoDynamics (QCD).

The weak force is the cause of radioactive decays, such as β-decay. It allows the coupling of different flavours of fermions, causing the decay of particles that cannot decay via the other forces due to conservation laws. It also allows Charge-Parity (*CP*) violation to occur, which will be discussed later. There have been several theories describing the weak force in isolation. However, the current Standard Model uses a unification of the electromagnetic and weak forces called the ElectroWeak (EW) interaction. This is similar to the unification of the electric and magnetic forces into the electromagnetic force by Maxwell [7].

Gravity is the fourth fundamental force of nature and is the cause of the force of attraction between all particles with mass or energy. It is is well described in classical physics by Einstein's General Theory of Relativity [8]. The theoretical massless spin-2 particle called the graviton is the expected boson for the gravita-

tional force in a QFT of gravity. So far, a predictive QFT of gravity has not been formulated due to it being a fundamentally *non-renormalizable* theory [9]. This means that it has infinitely many independent free parameters which would have to be measured in order to describe the theory and is therefore currently useless as a theory of nature.

Form of the Standard Model

In the Standard Model, the dynamics of the particles and particle interactions are described using a Lagrangian density ($\hat{\mathcal{L}}$). This Lagrangian has many different terms describing such things as the propagation of a free particle, the coupling of particles to different fields, the self interactions of fields and the acquisition of mass. The symmetry properties of this Lagrangian are intimately tied to both the characteristics of the fundamental interactions it describes and the conservation of quantum numbers associated with those interactions.

The relationship between continuous symmetries and conservation laws in a classical system is described by Noether's theorem [10]. This states that if the equations of motion derived from a Lagrangian are invariant when a continuous transformation is applied, then there will be a conserved quantity. For example, it can be shown that the invariance of physical laws to position are tied to the conservation of the classical momentum. Noether's theorem does extend to QFTs and the result is that invariances of a Lagrangian cause conserved *symmetry currents*, which correspond to physical properties conserved in particle interactions, such as electric charge. It is therefore beneficial to use symmetries and Lagrangians when building a theory of nature, as the observed conservation laws can be built into the theory by making it observe specific symmetries.

The Standard Model is a gauge theory, meaning that the Lagrangian is invariant under certain continuous local transformations, or symmetries. As an example, the gauge freedom of the Standard Model can be understood from the QED Lagrangian density

$$\hat{\mathcal{L}}_{\mathrm{QED}} = \hat{\mathcal{L}}_{\mathrm{D}} + \hat{\mathcal{L}}_{\mathrm{Int}} + \hat{\mathcal{L}}_{\mathrm{Em}} \tag{1.1}$$

for a fermion of mass, m, and charge, q, where

$$\hat{\mathcal{L}}_{\mathrm{D}} = \hat{\bar{\psi}}(i\gamma^{\mu}\partial_{\mu} - m)\hat{\psi} \tag{1.2}$$

$$\hat{\mathcal{L}}_{\mathrm{Int}} = -q\hat{\bar{\psi}}\gamma^{\mu}\hat{\psi}\hat{A}_{\mu} \tag{1.3}$$

$$\hat{\mathcal{L}}_{\mathrm{Em}} = -\frac{1}{4}\hat{F}_{\mu\nu}\hat{F}^{\mu\nu} \tag{1.4}$$

$$\hat{F}_{\mu\nu} = \partial_{\mu}\hat{A}_{\nu} + \partial_{\nu}\hat{A}_{\mu}. \tag{1.5}$$

Here $\hat{\psi}$ is a quantised Dirac spinor field with γ^{μ} being the Dirac matrices. The term $\hat{A}_{\mu}$ is the covariant electromagnetic four-potential and $\hat{F}_{\mu\nu}$ is the Faraday tensor.

The three Lagrangian terms that make up the QED Lagrangian describe different dynamics of QED. The term $\hat{\mathcal{L}}_{\mathrm{D}}$ controls how fermions interact with one another, while $\hat{\mathcal{L}}_{\mathrm{Int}}$ describes the interaction of charged fermions with photon fields. The final term $\hat{\mathcal{L}}_{\mathrm{Em}}$ explains the propagation of photons in free space without any interaction with fermions, as shown by the lack of Dirac spinors in this term. The total QED Lagrangian is invariant under local U(1) transformations *i.e.* transformations of the form

$$\hat{\psi} \mapsto \hat{\psi}' = e^{-iq\hat{\chi}(x^{\mu})}\hat{\psi} \tag{1.6}$$

Here $\hat{\psi}$ is a quantised Dirac spinor field, $\hat{\chi}(x^{\mu})$ is any general quantised scalar field and $\hat{A}_{\mu}$ is the covariant electromagnetic four-potential. The invariance of the QED Lagrangian to this transformation leads to the conservation of the electric charge, q, and the introduction of $\hat{A}_{\mu}$ corresponds to a single vector gauge boson, the photon.

The full Standard Model Lagrangian is locally invariant under the unitary $\mathrm{SU}(3)\times\mathrm{SU}(2)\times\mathrm{U}(1)$ gauge group. The SU(3) comes from QCD and the $\mathrm{SU}(2)\times\mathrm{U}(1)$ from the EW terms[I].

The electroweak theory

In 1967 the EW theory was described through the combined efforts of Glashow, Weinberg and Salam [11–13]. As discussed earlier, it was a successful attempt to make a unified gauge theory for the weak and electromagnetic interactions using the $\mathrm{SU}(2) \times \mathrm{U}(1)$ gauge group.

An important feature of the theory is the inherent *chirality* of the weak interaction. Chirality implies that the left and right-handed components of a wave-function interact differently with the weak interaction. To understand what is meant

[I]More precisely the notation should be $\mathrm{SU}(2)_T \times \mathrm{U}(1)_Y$ indicating the relations of the groups to the weak isospin, T, and the weak hyper-charge, Y. These will be omitted unless otherwise stated.

by left and right-handed components, first consider the general Dirac component of the Lagrangian in Eq. 1.2. It is possible to split a Dirac spinor field, $\hat{\psi}$, into the left and right chiral components

$$\hat{\psi} = \hat{\psi}_L + \hat{\psi}_R, \tag{1.7}$$

where

$$\hat{\psi}_R = \frac{1}{2}(1+\gamma_5)\hat{\psi} \quad \text{and} \quad \hat{\psi}_L = \frac{1}{2}(1-\gamma_5)\hat{\psi} \tag{1.8}$$

with similar terms for the adjoint spinor fields.

$$\hat{\bar{\psi}}_R = \hat{\bar{\psi}}\frac{1}{2}(1-\gamma_5) \quad \text{and} \quad \hat{\bar{\psi}}_L = \hat{\bar{\psi}}\frac{1}{2}(1+\gamma_5) \tag{1.9}$$

These decompositions use the fifth Dirac gamma matrix, $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$. Eq. 1.2 can be rewritten as

$$\hat{\mathcal{L}}_{\mathrm{D}} = \hat{\bar{\psi}}_R i\gamma^\mu\partial_\mu\hat{\psi}_R + \hat{\bar{\psi}}_L i\gamma^\mu\partial_\mu\hat{\psi}_L - m(\hat{\bar{\psi}}_R\hat{\psi}_L + \hat{\bar{\psi}}_L\hat{\psi}_R). \tag{1.10}$$

Here it is obvious that the left and right spinor fields interact separately for the kinetic portion of the Lagrangian, while the mass terms involve coupling of the left and right-handed components. It is therefore possible to treat the left and right-handed components as separate fields which can couple in different ways in the EW theory. It was observed experimentally that the charged current (CC) weak interactions violate parity, as discussed in Section 1.4. This implies that the CC weak interaction couples only to the left-handed spinor fields, while the neutral current (NC) can couple to both the left and right-handed components. Therefore the structure of the EW theory described below reflects this.

In order to form a concise representation of the symmetry groups, the fields for related particles are combined into multiplets. The left-handed leptons are formed into SU(2) *weak isospin* doublets

$$\hat{L}_{\mathrm{e}} = \begin{pmatrix} \hat{\nu}_e \\ \hat{\mathrm{e}} \end{pmatrix}_L \quad \hat{L}_{\mu} = \begin{pmatrix} \hat{\nu}_\mu \\ \hat{\mu} \end{pmatrix}_L \quad \hat{L}_{\tau} = \begin{pmatrix} \hat{\nu}_\tau \\ \hat{\tau} \end{pmatrix}_L, \tag{1.11}$$

while the right-handed charged[II] leptons form into the *weak iso-singlets*

$$\hat{R}_{\mathrm{e}} = \hat{\mathrm{e}}_R \quad \hat{R}_{\mu} = \hat{\mu}_R \quad \hat{R}_{\tau} = \hat{\tau}_R. \tag{1.12}$$

[II] In this simple treatment, the neutrinos are taken to be massless and naturally left-handed. There are therefore no right-handed neutrino singlets.

Here the chirality of the weak interaction is coded into the theory by the different representations of the chiral components of the fields. The quarks also form left-handed doublets and right-handed singlets, however the doublets are complicated by the difference between mass and weak eigenstates. This will be discussed in Section 1.3.

This representation of the fields allows the kinetic terms of a Lagrangian for the EW interaction with leptons to be written as

$$\hat{\mathcal{L}} = \sum_{\mathrm{f=e,\mu,\tau}} \hat{\bar{L}}_{\mathrm{f}} i\gamma^{\mu}\partial_{\mu}\hat{L}_{\mathrm{f}} + \sum_{\mathrm{f=e,\mu,\tau}} \hat{\bar{R}}_{\mathrm{f}} i\gamma^{\mu}\partial_{\mu}\hat{R}_{\mathrm{f}} \tag{1.13}$$

If covariant derivatives, $\hat{D}_{\mu}$, are introduced this Lagrangian can be shown to be invariant with respect to a $\mathrm{SU}(2)_T \times \mathrm{U}(1)_Y$ group gauge transformation. Namely one of the form

$$\hat{\psi} \mapsto \hat{\psi}' = e^{\frac{i}{2}\hat{\boldsymbol{\alpha}}(x^{\mu}).\boldsymbol{\tau} + i\hat{\beta}(x^{\mu})Y}\hat{\psi}, \tag{1.14}$$

where $\boldsymbol{\tau}$ is the vector of Pauli matrices. Here, $\hat{\boldsymbol{\alpha}}(x^{\mu})$ and $\hat{\beta}(x^{\mu})$ are a vector of general quantised fields and a general quantised field respectively. Also Y is the weak hyper-charge, given by its relation with the electric charge, Q, and the third component of the weak-isospin T_3

$$Q = T_3 + \frac{1}{2}Y. \tag{1.15}$$

The fact that the right-handed components do not transform under the SU(2) weak isospin group means that they do not carry the weak isospin quantum number and cannot couple to the CC weak interaction.

The covariant derivative has the form,

$$\hat{D}_{\mu} = \partial_{\mu} + \frac{ig}{2}\boldsymbol{\tau}\cdot\hat{\boldsymbol{W}}_{\mu} + \frac{ig'Y}{2}\hat{B}_{\mu}. \tag{1.16}$$

Here the coefficients g and g' are the coupling strengths of the SU(2) and U(1) interactions respectively. It should be noted that as the right-handed spinors do not couple to the weak force, the term $\frac{ig}{2}\boldsymbol{\tau}\cdot\hat{\boldsymbol{W}}_{\mu}$ is zero for the right-handed components.

$\hat{D}_{\mu}$ can be substituted into Eq. 1.13 for ∂_{μ} giving the gauge invariant Lagrangian

$$\hat{\mathcal{L}} = \sum_{\mathrm{f=e,\mu,\tau}} \hat{\bar{L}}_{\mathrm{f}} i\gamma^{\mu}\hat{D}_{\mu}\hat{L}_{\mathrm{f}} + \sum_{\mathrm{f=e,\mu,\tau}} \hat{\bar{R}}_{\mathrm{f}} i\gamma^{\mu}\hat{D}_{\mu}\hat{R}_{\mathrm{f}} \tag{1.17}$$

The EW interaction Lagrangian in Eq. 1.17 can also contain a gauge boson

term describing the propagation and interaction of the gauge boson fields in a similar way to Eq. 1.4. This term is invariant with respect to the transformation in Eq. 1.14 and has the form

$$\hat{\mathcal{L}}_{\text{Gauge}} = -\frac{1}{4}\hat{\boldsymbol{F}}_{\mu\nu}\cdot\hat{\boldsymbol{F}}^{\mu\nu} - \frac{1}{4}\hat{G}_{\mu\nu}\hat{G}^{\mu\nu} \tag{1.18}$$

$$\hat{\boldsymbol{F}}_{\mu\nu} = \partial_\mu\hat{\boldsymbol{W}}_\nu - \partial_\nu\hat{\boldsymbol{W}}_\mu - g\hat{\boldsymbol{W}}_\mu\times\hat{\boldsymbol{W}}_\nu \tag{1.19}$$

$$\hat{G}_{\mu\nu} = \partial_\mu\hat{B}_\nu - \partial_\nu\hat{B}_\mu, \tag{1.20}$$

The interaction Lagrangian in Eq. 1.18 contains four massless gauge bosons, $W_\mu^{1,2,3}$ and B_μ. Obviously these gauge bosons are not the physical ones listed in Table 1.1 since the Standard Model has only one massless gauge boson for the EW interaction (the photon), while the $\mathrm{W}^\pm$ and Z^0 have large masses. Unfortunately the EW interaction described so far cannot include masses, as the mass terms in the Lagrangians involve the coupling of the left and right-handed components of the fields, as shown in Eq. 1.10. Since the left and right-handed components are described by different multiplet representations there is no way to have standard mass terms without breaking the symmetry of the EW gauge group.

Spontaneous symmetry breaking and massive vector bosons

In order for the EW theory to have the correct vector bosons, it is necessary for the Lagrangian to somehow acquire mass terms without breaking the gauge invariance. This is achieved by the spontaneous symmetry breaking mechanism, which causes the full Lagrangian to remain invariant, while the vacuum state does not. In the case of the Standard Model, the weak vector bosons should acquire masses while the photon should remain massless. The simplest way of doing this is to introduce a weak isospin complex scalar doublet field called the Higgs field containing a charged and neutral field

$$\hat{\phi} = \begin{pmatrix} \hat{\phi}^+ \\ \hat{\phi}^0 \end{pmatrix}. \tag{1.21}$$

The Higgs field adds Lagrangian terms[III] in $\hat{\phi}$ which can be parameterised as

$$\hat{\mathcal{L}}_{\text{Higgs}} = (\hat{D}_\mu\hat{\phi})^\dagger(\hat{D}^\mu\hat{\phi}) + \mu^2\hat{\phi}^\dagger\hat{\phi} - \frac{\lambda}{4}(\hat{\phi}^\dagger\hat{\phi})^2. \tag{1.22}$$

[III]The terms can only be up to quartic in $\hat{\phi}$ as higher terms would make the Lagrangian non-renormalisable

Similar covariant derivatives to those in Eq. 1.16 are used, and μ and λ are constants.

It is necessary to have a non-zero Vacuum Expectation Value (VEV), v, of this field and the choice suggested by Weinberg in 1967 was

$$\langle 0|\hat{\phi}|0\rangle = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} \tag{1.23}$$

where $\pm\frac{v}{\sqrt{2}} = \sqrt{\frac{2\mu^2}{\lambda}}$. This choice prevents the appearance of a massive photon by setting the charged field VEV to zero.

By considering perturbations around the vacuum state, then adding and expanding the Lagrangians from Eq. 1.18 and Eq. 1.22, it is possible to find quadratic terms in the Lagrangian and make the assignments

$$\hat{Z}_\mu = \cos\theta_{\mathrm{W}}\hat{W}^3_\mu - \sin\theta_{\mathrm{W}}\hat{B}_\mu \tag{1.24}$$

$$\hat{A}_\mu = \sin\theta_{\mathrm{W}}\hat{W}^3_\mu + \cos\theta_{\mathrm{W}}\hat{B}_\mu, \tag{1.25}$$

with

$$\sin\theta_{\mathrm{W}} = g/(g^2+g'^2)^{\frac{1}{2}} \tag{1.26}$$

$$\cos\theta_{\mathrm{W}} = g'/(g^2+g'^2)^{\frac{1}{2}}. \tag{1.27}$$

Here θ_{W} is the Weinberg angle. The fields $\hat{Z}_\mu$ and $\hat{A}_\mu$ are the massive Z^0 and photon vector boson fields respectively. The photon field is combined into the QED gauge field term, *cf.* Eqs. 1.4-1.5, which contains no mass term for the photon as expected. Also, due to the scalar Higgs field, there is a mass term for the physical Higgs boson

$$m_{\mathrm{H}} = \sqrt{2}\mu = \sqrt{2\lambda}v. \tag{1.28}$$

Using the coefficients of Eq. 1.22 and θ_{W}, the masses of the vector bosons can finally be calculated to be

$$m_{\mathrm{W}} = \frac{gv}{2} \tag{1.29}$$

$$m_{\mathrm{Z}} = \frac{m_{\mathrm{W}}}{\cos\theta_{\mathrm{W}}}. \tag{1.30}$$

1.3 Quark mixing and the CKM matrix

The previous discussion of the EW theory focused on the appearance of gauge boson masses in the Standard Model. However, the fermions are also observed to be massive and the theory should reflect this. The following section will introduce the Lagrangian terms responsible for fermion masses and the consequences in the quark sector.

Yukawa coupling and quark masses

The chirality of the weak interaction again prevents the introduction of Dirac mass terms for the fermion fields of the form Eq. 1.10, without explicitly breaking the gauge symmetry. However, since the EW theory already has spontaneous symmetry breaking for the gauge boson masses it is possible to give the fermions masses another way. Take the lepton singlets and doublets, Eq. 1.12, Eq. 1.11, and the Higgs doublet, Eq. 1.23, as a starting point. If the leptons have Yukawa type coupling to the Higgs field

$$\hat{\mathcal{L}}_{\mathrm{Y}}^{l} = \sum_{\mathrm{f=e,\mu,\tau}} -y_{\mathrm{f}}(\hat{\bar{L}}_{\mathrm{f}}\hat{\phi}\hat{R}_{\mathrm{f}} + \hat{\bar{R}}_{\mathrm{f}}\hat{\phi}^{\dagger}\hat{L}_{\mathrm{f}}), \tag{1.31}$$

lepton mass terms will be generated with $m_{\mathrm{f}} = y_{\mathrm{f}}v/\sqrt{2}$, so long as y_{f} are real and positive [14].

To define the Yukawa terms for the quark fields, first the quark doublets

$$\hat{Q}_i^L = \begin{pmatrix} \hat{u}_i^L \\ \hat{d}_i^L \end{pmatrix} \quad \text{for} \quad i = 1,2,3 \tag{1.32}$$

and singlets

$$\hat{u}_i^R, \hat{d}_i^R \quad \text{for} \quad i = 1,2,3. \tag{1.33}$$

are defined. The Yukawa terms for the quark fields can then be generalised as

$$\hat{\mathcal{L}}_{\mathrm{Yuk}}^{q} = -\sum_{i,j=1}^{3} \left(a_{ij}\hat{\bar{Q}}_i^L \hat{\phi}^{\boldsymbol{c}} \hat{u}_j^R + b_{ij}\hat{\bar{Q}}_i^L \hat{\phi}\hat{d}_j^R + h.c. \right), \tag{1.34}$$

where $h.c.$ stands for the Hermitian conjugate terms and $\hat{\phi}^{\boldsymbol{c}}$ is the charge conjugate of the Higgs doublet. The matrix elements a_{ij} and b_{ij} give the mass matrices $M_u \equiv m_{ij}^u = va_{ij}/\sqrt{2}$ and $M_d \equiv m_{ij}^d = vb_{ij}/\sqrt{2}$. These will correspond to the physical quark masses if they are diagonal and Hermitian, however in this general

case there is no such requirement [15].

In order to diagonalise the mass matrices a set of unitary matrix transformations can be applied

$$V_{dL} M_d V_{dR}^{\dagger} = M_d^{\text{diag}} \tag{1.35}$$
$$V_{uL} M_u V_{uR}^{\dagger} = M_u^{\text{diag}} \tag{1.36}$$

allowing the unitary mixing matrix $V = V_{uL} V_{dL}^{\dagger}$ to be defined [4, 16].

The CKM matrix

The mixing matrix V is known as the CKM matrix, V_{CKM}. Applying V_{CKM} to the mass eigenstates triplet of the down-type quarks creates the weak eigenstate triplet

$$\begin{pmatrix} \hat{\hat{d}}'_L \\ \hat{\hat{s}}'_L \\ \hat{\hat{b}}'_L \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} \hat{\hat{d}}_L \\ \hat{\hat{s}}_L \\ \hat{\hat{b}}_L \end{pmatrix}, \tag{1.37}$$

which the weak CC interactions couple to. These interaction terms take the form

$$\hat{\mathcal{L}}_{\text{CC}} = -\frac{g}{\sqrt{2}} (\hat{\hat{u}}_L, \hat{\hat{c}}_L, \hat{\hat{t}}_L) \gamma^{\mu} V_{\text{CKM}} \begin{pmatrix} \hat{\hat{d}}_L \\ \hat{\hat{s}}_L \\ \hat{\hat{b}}_L \end{pmatrix} W_{\mu}^{\dagger} + h.c. \tag{1.38}$$

Since the CKM matrix is a 3×3 matrix[IV] it appears to have nine free complex terms. However, since the matrix is unitary there are further constraints which reduce its degrees of freedom. Five of the potential degrees of freedom are unobservable and they can be absorbed into the quark fields. This leaves four degrees of freedom in the CKM matrix. It can be shown that the CKM matrix can be parameterised by three mixing parameters and one non-trivial phase. This makes the matrix potentially complex if the imaginary component is non-zero and causes the weak CC interactions to violate *CP* symmetry, as discussed later. The constraint of unitarity also prevents Flavour Changing Neutral Currents (FCNCs) from appearing in the theory. This means that the Z^0 boson cannot couple to differently flavoured same-charge particles.

The standard parameterisation of the CKM matrix is [17–19]

[IV] The matrix is actually an $n \times n$ unitary matrix where n is the number of generations. The number of degrees of freedom increases with the number of generations.

$$V_{\rm CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \quad (1.39)$$

where $c_{ij} = \cos\theta_{ij}, s_{ij} = \sin\theta_{ij}$. The three real parameters are θ_{ij} and δ is the remaining phase.

The magnitudes of the elements of $V_{\rm CKM}$ are responsible for the relative coupling strengths between different flavours of quarks. These magnitudes have been experimentally measured [4] to be

$$\begin{aligned} |V_{\rm CKM}| &= \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} \\ &= \begin{pmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.0412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{pmatrix}. \end{aligned} \quad (1.40)$$

The Wolfenstein parameterisation

Alternative parameterisations have been proposed for the CKM matrix, often relating to the experimentally observed values of the CKM parameters and the hierarchy of the magnitudes of the elements seen in Eq. 1.40. An often used form is the Wolfenstein parameterisation [20]

$$V_{\rm CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \quad (1.41)$$

which uses the order of the λ expansion parameter to show the explicit hierarchy of the matrix elements, and η is the imaginary part. This parametrisation has the benefit of all four parameters being $\mathcal{O}(1)$. Additionally the following conversions between the standard and Wolfenstein parameterisations can be made

$$s_{12} = \lambda \tag{1.42}$$
$$s_{23} = A\lambda^2 \tag{1.43}$$
$$s_{13}e^{i\delta} = A\lambda^3(\rho + i\eta). \tag{1.44}$$

1.4 *CP*-violation

Discrete symmetries

The continuous symmetries described in Section 1.2 are not the only symmetries of nature. There are also three discrete symmetries of the Standard Model, Parity (P), Time (T) and Charge conjugation (C). Transformations of P and T are space-time transformations which cause $x^\mu \mapsto x_\mu$ and $x^\mu \mapsto -x_\mu$ respectively. The transformation C causes a particle to be transformed into its corresponding antiparticle by changing the internal quantum numbers *e.g.* electromagnetic charge.

Any general local QFT which is invariant under Poincaré group (Lorentz) transformations will be invariant with respect to the product of all three discrete transformations, *CPT* [14, 21]. As a consequence of this invariance, fundamental particles and their antiparticles have the same masses and total decay widths.

A history of *CP*-violation

Both the strong and electromagnetic force are invariant with respect to parity. The fact that violation of parity in weak interactions had not been experimentally tested was first proposed by Yang and Lee in 1956 [22]. After the suggestions of possible tests by Yang and Lee, parity violation was then observed in 1957 by Wu [23] through the decay of ^{60}Co. Parity violation in weak interactions is actually 'maximal' due to the weak interaction only coupling to the left-handed components of the spinor fields, rather than a mix of left and right-handed fields *cf.* Section 1.2.

This behaviour of the weak interaction seemed extremely strange at the time. It was proposed that parity was not a true symmetry of nature, but rather that the combined operation, *CP*, was. Since *CP* transforms a left-handed particle into a right-handed antiparticle, this would mean that the equivalence of matter and antimatter decays was the correct symmetry of nature.

Applying a *CP* transformation to a weak coupling such as Eq. 1.38 causes it to become its Hermitian conjugate. It is then apparent that the weak couplings are only invariant with respect to *CP* if $V^{ij}_{\mathrm{CKM}} = V^{ij*}_{\mathrm{CKM}}$ after all trivial phases are

removed. As already discussed in Section 1.3 the CKM matrix does have a non-trivial phase, leading to the possibility of observable *CP*-violation in certain quark transitions. *CP*-violation was first observed in the decays of neutral kaons in 1964 by Cronin and Fitch [24] and first observed in neutral B-meson system at BABAR and BELLE in 2001 [25, 26].[V]

Unitarity triangles

The fact that the CKM matrix must be unitary to satisfy Eqs. 1.35-1.36 means that there are 12 orthogonality constraints placed upon the elements

$$\sum_{i=\mathrm{u,c,t}}^{n} V_{ij}V_{ik}^{*} = \delta_{jk} \tag{1.45}$$

$$\sum_{j=\mathrm{d,s,b}}^{n} V_{ij}V_{kj}^{*} = \delta_{ik}, \tag{1.46}$$

where n is the number of quark generations. For the Standard Model CKM matrix, the most commonly used equation is

$$V_{ud}V_{ub}^{*} + V_{cd}V_{cb}^{*} + V_{td}V_{tb}^{*} = 0, \tag{1.47}$$

due to each of the terms having equal order in the Wolfenstein parameterisation. This equation describes a triangle in the complex plane, called the 'Unitarity triangle'.

It is useful to describe this triangle using relations found from the Wolfenstein parameters [28]. Firstly, the best measured term in this equation can be written as

$$V_{cd}V_{cb}^{*} = -A\lambda^{3} + \mathcal{O}(\lambda^{7}). \tag{1.48}$$

This means that to a high degree of accuracy $V_{cd}V_{cb}^{*}$ is real and $|V_{cd}V_{cb}^{*}| = A\lambda^{3}$. By rescaling Eq. 1.47 by $|V_{cd}V_{cb}^{*}|$ and keeping $\mathcal{O}(\lambda^{5})$ corrections, the relations

$$\frac{V_{ud}V_{ub}^{*}}{|V_{cd}V_{cb}^{*}|} = \overline{\rho} + i\overline{\eta} \tag{1.49}$$

$$\frac{V_{td}V_{tb}^{*}}{|V_{cd}V_{cb}^{*}|} = 1 - (\overline{\rho} + i\overline{\eta}) \tag{1.50}$$

[V] Due to the invariance of the Standard Model Lagrangian to *CPT* transformations, the violation of *CP* implies that T is also violated. However, T-violation can be searched for independently of *CP*-violation and has recently been observed by BABAR [27].

can be found. The parameters $\overline{\rho}$ and $\overline{\eta}$ are related to the Wolfenstein parameters ρ and η by the relations

$$\overline{\rho} \simeq \rho(1 - \frac{\lambda^2}{2}) \quad \text{and} \quad \overline{\eta} \simeq \eta(1 - \frac{\lambda^2}{2}), \tag{1.51}$$

which are correct to $\mathcal{O}(\lambda^2)$. Fig. 1.2 shows that by plotting the rescaled triangle on the complex plane, one side is of unit length and lies on the real axis and the top vertex corresponds to the coordinates $(\overline{\rho}, \overline{\eta})$. This rescaling and re-parameterisation allows for many measurements to be easily visualised on this plane, see Fig. 1.3.

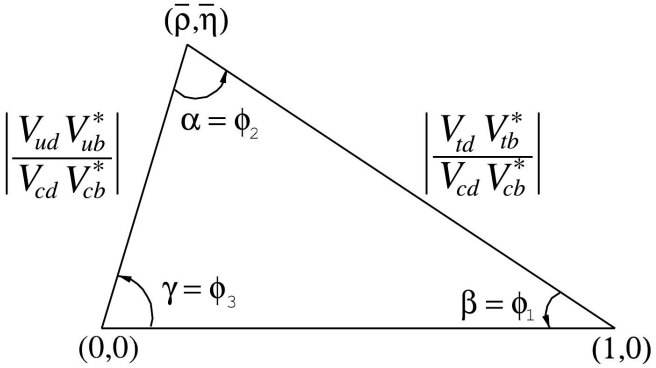

Figure 1.2: A sketch of the unitarity triangle [4]. The notation for the angles α, β, γ is used by BABAR and LHCb. BELLE uses the alternate notation ϕ_1, ϕ_2, ϕ_3.

Additionally, the internal angles of the unitarity triangle are also intimately related to the elements of the CKM matrix. Some commonly used relations are

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right) \tag{1.52}$$

$$\beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) \tag{1.53}$$

$$\gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \tag{1.54}$$

The Jarlskog invariant, J, is a phase-convention independent measure of the amount of *CP*-violation in the Standard Model [29]. It is equal to twice the area of the unitary triangles and is given by the equation

$$\mathcal{I}m(V_{ij}V_{kl}V_{il}^*V_{kj}^*) = J\sum_{m,n}\epsilon_{ikm}\epsilon_{jln}. \tag{1.55}$$

The measured value of the Jarlskog invariant is, $J = (2.96^{+0.20}_{-0.16}) \times 10^{-5}$ [4].

Measuring the sides and angles of the unitarity triangle is a sensitive method for testing the Standard Model. For example, if there are more than three generations of quarks, the unitarity triangle will not close. Therefore if the unitarity triangle can be over-constrained by precision measurements, the existence of physics beyond the Standard Model can be determined. The current measured constraints on the sides and angles of the unitarity angle are shown in Fig. 1.3.

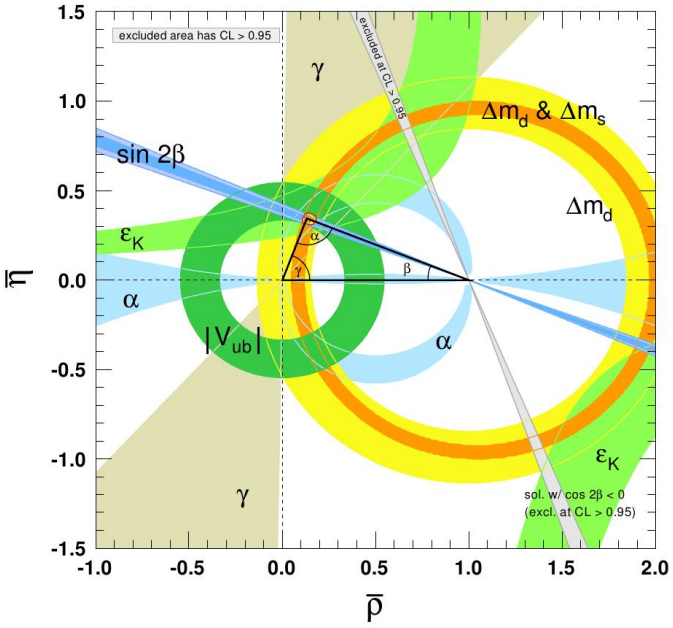

Figure 1.3: Constraints on the angles and sides of the unitarity triangle in the $\overline{\rho}$-$\overline{\eta}$ plane with 95% confidence limits [4].

As quarks are only found inside bound states of baryons and mesons it is necessary to perform the measurements of the unitarity triangle properties from the decays of the these composite particles. Therefore, the theoretical framework of flavour physics usually describes the interactions of the measurable hadrons, rather than the confined quarks. There are several mechanisms by which measurable *CP*-violation can occur for the decay of a hadron, specifically for neutral B mesons which are the focus of this analysis.

Direct *CP*-violation

When the amplitude for a decay is not equal to the amplitude for its *CP* conjugate decay, *CP*-violation is said to have occurred directly in the decay. For the decay

of a B to a final *CP* eigenstate f, $B \to f$, and the *CP* conjugate decay, $\overline{B} \to \overline{f}$, the amplitudes for these decays can be written as

$$A_f = \sum_k A_k e^{i(\delta_k + \phi_k)}, \quad \bar{A}_{\overline{f}} = \sum_k A_k e^{i(\delta_k - \phi_k)}, \tag{1.56}$$

where the index, k indicates a contributing decay process. A_k is the magnitude of the amplitude, δ_k and ϕ_k are the *CP*-even and *CP*-odd components of the phase for the process.

In order for *CP*-violation to occur, $|\bar{A}_{\overline{f}}/A_f| \neq 1$. This can only occur if at least two terms have *CP*-odd and *CP*-even phases which are different from each other. Direct *CP*-violation was first observed in the decays of kaons [30, 31] and subsequently in B meson decays at BABAR [32].

CP-violation in mixing

As discussed in Section 1.3, the mass and flavour eigenstates of quarks are not necessarily the same. For neutral mesons, such as B^0, there are the flavour eigenstates of the B^0 and $\overline{B}^0$ denoted by $|B^0\rangle$ and $|\overline{B}^0\rangle$. These states can undergo box diagram mixing transitions like those in Fig. 1.4, producing a superposition of the B^0 and $\overline{B}^0$ states.

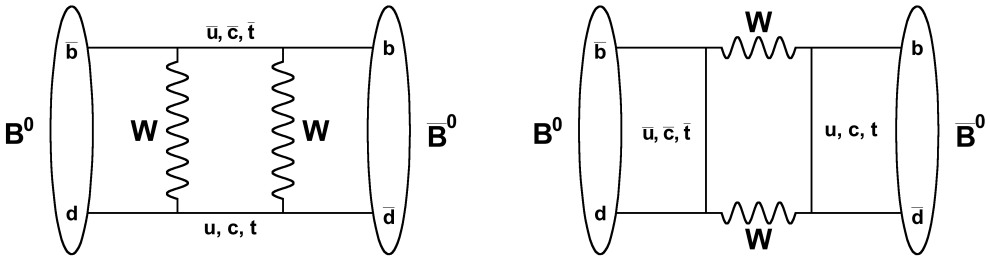

Figure 1.4: Feynman diagrams of B^0-$\overline{B}^0$ mixing.

However, the time evolution of the superposition is governed by the time-dependent Schrödinger equation

$$i\frac{d}{dt}\begin{pmatrix} |B^0(t)\rangle \\ |\overline{B}^0(t)\rangle \end{pmatrix} = \left(M - \frac{i}{2}\Gamma\right)\begin{pmatrix} |B^0(t)\rangle \\ |\overline{B}^0(t)\rangle \end{pmatrix}, \tag{1.57}$$

where M and Γ are independent Hermitian matrices and non-zero off diagonal elements imply that the flavour and mass eigenstates are different. The mass eigen-

states of the B states are the eigenvectors of the Hamiltonian, $H = M - \frac{i}{2}\Gamma$, and can be written as linear superpositions of the flavour eigenstates

$$\begin{aligned} \text{Light eigenstate}: \quad |\mathrm{B_L}\rangle &= p|\mathrm{B^0}\rangle + q|\overline{\mathrm{B}}^0\rangle \\ \text{Heavy eigenstate}: \quad |\mathrm{B_H}\rangle &= p|\mathrm{B^0}\rangle - q|\overline{\mathrm{B}}^0\rangle. \end{aligned} \tag{1.58}$$

Here p and q are complex coefficients with $|p|^2 + |q|^2 = 1$. *CP* symmetry requires that $H_{12}^* = H_{21}$, leading to the requirement that for *CP*-violation due to mixing to occur, $|\frac{p}{q}| \neq 1$.

The evolution of the mass eigenstates is governed by the eigenvalues $M_\mathrm{H} - \frac{i}{2}\Gamma_\mathrm{H}$ and $M_\mathrm{L} - \frac{i}{2}\Gamma_\mathrm{L}$. The eigenvalues also allow the quantities $\Delta m = M_\mathrm{H} - M_\mathrm{L}$ and $\Delta\Gamma = \Gamma_\mathrm{H} - \Gamma_\mathrm{L}$ to be defined. These are the mass and decay width splittings of the mass eigenstates.

Solving for the eigenvalues and eigenvectors of H gives

$$\left(\frac{q}{p}\right)^2 = \frac{M_{12}^* - (i/2)\Gamma_{12}^*}{M_{12} - (i/2)\Gamma_{12}}. \tag{1.59}$$

As it is experimentally found that $\Delta\Gamma << \Delta m$ in both the $\mathrm{B^0}$ and $\mathrm{B_s^0}$ systems, a leading order estimate of $|\frac{p}{q}|$ gives unity, implying that mixing in this sector is small. From the upper bounds on the *CP* asymmetry in semileptonic B decays, it is found that for the $\mathrm{B^0}$ system that $|q/p| = 1.0017 \pm 0.0017$ and in the $\mathrm{B_s^0}$ system $|q/p| = 1.0052 \pm 0.0032$, confirming this expectation [4].

CP-violation in interference between mixing and decay

The last form of *CP*-violation occurs when both a particle and its antiparticle decay to the same *CP* eigenstate f ($\mathrm{f} = \overline{\mathrm{f}}$). The two decay processes $\mathrm{B} \to \mathrm{f}$ and $\mathrm{B} \to \overline{\mathrm{B}} \to \mathrm{f}$ can cause *CP*-violation if there is interference between the *CP*-violation in mixing and the *CP*-violation in the decay. To quantify the interference, the parameter λ_f is given by

$$\lambda_\mathrm{f} = \frac{q}{p}\frac{\bar{A}_\mathrm{f}}{A_\mathrm{f}}, \tag{1.60}$$

which is equal to unity if there is no *CP*-violation from interference. It is possible that $|\lambda_\mathrm{f}| = 1$ when there is no *CP*-violation from the decay or mixing on their own, but that $\mathcal{I}m(\lambda_\mathrm{f}) \neq 0$ due to the interference of the two *CP*-conserving processes.

CP-violation due to interference is measurable from the asymmetry of neutral

meson decays to a *CP* eigenstate f

$$\mathcal{A}_f(t) \equiv \frac{d\Gamma/dt[\overline{\mathrm{B}}(t) \to f] - d\Gamma/dt[\mathrm{B}(t) \to f]}{d\Gamma/dt[\overline{\mathrm{B}}(t) \to f] + d\Gamma/dt[\mathrm{B}(t) \to f]}. \tag{1.61}$$

For B^0 meson decays the good approximations, $|q/p| = 1$ and $\Delta\Gamma = 0$, can be made. This allows Eq. 1.61 to be re-written as

$$\mathcal{A}_f(t) = S_f \sin(\Delta m t) - C_f \cos(\Delta m t), \tag{1.62}$$

with

$$S_f = \frac{2\mathcal{I}m(\lambda_f)}{1 + |\lambda_f|^2}, \quad C_f = \frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2}. \tag{1.63}$$

In the B_s^0 system, $|q/p| = 1$, is expected to hold to a good approximation. However $\Delta\Gamma_s$, the mass splitting of the B_s^0 system, has been measured at LHCb to be $0.106 \pm 0.011(\mathrm{stat.}) \pm 0.007\mathrm{syst.\,ps}^{-1}$, significantly different from zero [33]. In this case, the asymmetry of Eq. 1.61 for B_s^0 mesons becomes

$$\mathcal{A}_f(t) = \frac{\Gamma_{\overline{\mathrm{B}}_s^0}(t) - \Gamma_{\mathrm{B}_s^0}(t)}{\Gamma_{\overline{\mathrm{B}}_s^0}(t) + \Gamma_{\mathrm{B}_s^0}(t)} = \frac{-C_f \cos(\Delta m_s t) + S_f \sin(\Delta m_s t)}{\cosh\left(\frac{\Delta\Gamma_s}{2}t\right) - A_f^{\Delta\Gamma} \sinh\left(\frac{\Delta\Gamma_s}{2}t\right)}, \tag{1.64}$$

where S_f and C_f have the same meaning and

$$A_f^{\Delta\Gamma} = -\frac{2\mathcal{R}e(\lambda_f)}{1 + |\lambda_f|^2}. \tag{1.65}$$

Similar expressions hold for final states which are not *CP* eigenstates. Detailed descriptions can be found in [14, 34].

1.5 The $\mathrm{B}_s^0 \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$ decay mode

A penguin decay

Decay modes such as $\mathrm{B}_s^0 \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$, which do not proceed via a charm quark are known as 'charmless' modes. These decays will often proceed via 'penguin loop' decays, shown in Fig. 1.5, which allow the $\mathrm{b} \to \mathrm{s}$ quark transition via a loop. Although all three up-type quarks are possible in the loop, the top quark is dominant due to its much larger mass. Additionally, high mass undiscovered particles can potentially participate in the loop, leading to a different decay rate from the Standard Model

prediction. LHCb is designed specifically to perform the precision measurements required to test these types of predictions.

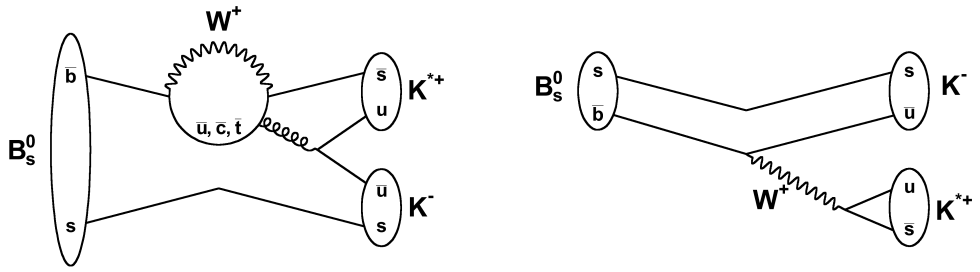

Figure 1.5: Feynman diagrams of the $B_s^0 \to K^{*+}K^-$ decay. Left: QCD penguin loop. Right: Tree diagram.

The corresponding tree-level diagram (without a loop) is suppressed by the CKM matrix elements. This means that, in general, the tree and penguin decay processes have similar magnitudes allowing the interference between the tree and penguin diagrams to be more easily determined.

The physics potential of the $B_s^0 \to K^{*\pm}K^{\mp}$ decay

Simple symmetry arguments can create useful tests of the Standard Model. U-spin symmetry is the equivalence of decays under the substitution of d and s quarks. Therefore comparison between B^0 and B_s^0 decay processes into final states containing a pion or kaon respectively become tests of this symmetry. Neglecting higher order annihilation terms, the relations

$$\mathcal{B}\left(B_s^0 \to K^{*+}K^-\right) \approx \mathcal{B}\left(B^0 \to \rho^+K^-\right) \quad (1.66)$$

$$\mathcal{B}\left(B_s^0 \to K^{*-}K^+\right) \approx \mathcal{B}\left(B^0 \to K^{*-}\pi^+\right), \quad (1.67)$$

can be found [35]. Therefore precise measurement of these branching fractions allows estimation of the contribution from the higher order processes.

As mentioned in Section 1.4, the sides and angles of the unitarity triangle are sensitive to physics beyond the Standard Model. Therefore, the precise determination of these values is an excellent method for indirectly observing the effects of new physics. It is well known that charmless modes are a rich area for the determination of the unitarity triangle's properties [36], specifically for the measurement of the angle γ. In the decays of $B^0 \to K^*\pi$ to the Isospin state, $I = 3/2$, the total

amplitude is given by

$$3A_{3/2} = A(\mathrm{B}^0 \to \mathrm{K}^{*+}\pi^-) + \sqrt{2}A(\mathrm{B}^0 \to \mathrm{K}^{*0}\pi^0), \tag{1.68}$$

with an equivalent equation for the *CP* conjugate process. This then allows γ to be measured, since neglecting ElectroWeak Penguin (EWP) diagrams leads to

$$\gamma = \Phi_{3/2} \equiv -\frac{1}{2}\arg\left(\frac{\bar{A}_{3/2}}{A_{3/2}}\right). \tag{1.69}$$

A sketch of the contributing amplitudes of Eq. 1.68 is shown in Fig. 1.6.

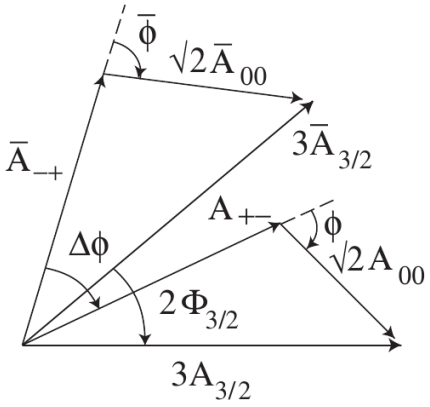

Figure 1.6: Sketch of the amplitudes for Eq. 1.68 and its *CP* conjugate, showing phase differences leading to Eq. 1.69 [36]

Similar arguments can be made in the B_s^0 system for the suppressed $\mathrm{B}_s^0 \to \mathrm{K}^*\pi$ decay modes [37] and for the favoured $\mathrm{B}_s^0 \to \mathrm{K}^*\mathrm{K}$ modes [38].

Previous measurements and theoretical predictions

The charmless 3-body decays of B mesons have been extensively studied and many of their branching fractions measured. In Table 1.2 some world averages and new measurements are presented for some of the charmless decays to similar final states.

The absence of a previous measurement for the $\mathrm{B}_s^0 \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$ decay modes is due to the running conditions of the B-factory experiments. The B-factories were run at the centre-of-mass energy of the $\Upsilon(4S)$ resonance, which decays predominately to a $\mathrm{B}\overline{\mathrm{B}}$ pair. This is excellent for the experimentally clean production of $\mathrm{B}\overline{\mathrm{B}}$ pairs, but the energy is too low to produce the B_s^0 mesons required for this analysis.

Table 1.2: Experimental measurements of charmless B decay branching fractions. † are the non-resonant components to the final state.

Measurement	Value ($\times 10^{-6}$)	Reference
$\mathcal{B}(\mathrm{B}^0 \to \mathrm{K}^0_\mathrm{S}\pi^+\pi^-)$	49.6 ± 2.0	[4]
$\mathcal{B}(\mathrm{B}^0 \to \mathrm{K}^0_\mathrm{S}\pi^+\pi^-)$ †	$14.7^{+4.0}_{-2.6}$	[4]
$\mathcal{B}(\mathrm{B}^0_s \to \mathrm{K}^0_\mathrm{S}\pi^+\pi^-)$	$11.9 \pm 3.0 \pm 2.1$	[1]
$\mathcal{B}(\mathrm{B}^0_s \to \mathrm{K}^0_\mathrm{S}\mathrm{K}^\pm\pi^\mp)$	$97 \pm 7 \pm 11$	[1]
$\mathcal{B}(\mathrm{B}^0_s \to \mathrm{K}^0_\mathrm{S}\mathrm{K}^+\mathrm{K}^-)$	$4.2 \pm 1.5 \pm 0.9$	[1]
$\mathcal{B}(\mathrm{B}^0 \to \mathrm{K}^{*+}\pi^-)$	8.5 ± 0.7	[39]

Previous measurements of $\mathrm{B} \to \mathrm{hh}$ decays have been made at a hadron collider by the CDF experiment [40], however no observations of the $\mathrm{B}^0_s \to \mathrm{K}^{*\pm}\mathrm{K}^\mp$ mode has yet been made. At the LHC B^0_s mesons are produced copiously at the LHCb interaction point which greatly increases the chance for observation, depending on the branching fraction values. Some theoretical predictions for the unmeasured $\mathrm{B}^0_{d,s} \to \mathrm{K}^{*\pm}\mathrm{h}^\mp$ decays are summarised in Table 1.3.

In order to use a single estimate of the branching fractions for yield estimates and easy comparison, weighted mean averages of the theoretical predictions were made for this analysis. The weights were taken as $w_i = 1/\sigma_i^2$ where σ_i is the uncertainty on the prediction i. The calculated averages for the relative branching fractions are given in Table 1.4.

Table 1.3: Theoretical predictions of unmeasured charmless $B^0_{d,s} \to K^{*\pm}h^{\mp}$ decay branching fractions. The three theoretical models are QCD Factorization (QCDF), Perturbative QCD (PQCD) and a six-quark Effective Hamiltonian method. Branching fractions given in units of ($\times 10^{-6}$). $K^{*\pm}K^{\mp}$ combined branching fractions and relative branching fractions are naive combinations made for this analysis. Uncertainties were added in quadrature and the experimentally measured value of $\mathcal{B}\left(B^0 \to K^{*+}\pi^-\right)$ was used to make the relative values.

Branching fraction	QCDF [35]	pQCD [41]	Eff. Hamiltonian [42]
$\mathcal{B}\left(B^0 \to K^{*+}K^-\right)$	$0.07^{+0.01+0.04}_{-0.01-0.03}$	-	-
$\mathcal{B}\left(B^0 \to K^{*-}K^+\right)$	$0.08^{+0.01+0.02}_{-0.01-0.02}$	-	-
$\mathcal{B}\left(B^0 \to K^{*\pm}K^{\mp}\right)$	0.15 ± 0.04	-	-
$\frac{\mathcal{B}\left(B^0 \to K^{*\pm}K^{\mp}\right)}{\mathcal{B}(B^0 \to K^{*+}\pi^-)}$	0.018 ± 0.005	-	-
$\mathcal{B}\left(B^0_s \to K^{*-}\pi^+\right)$	$7.8^{+0.4+0.5}_{-0.7-0.7}$	$7.6^{+2.9+0.4+0.5}_{-2.2-0.5-0.3}$	$7.2^{+5.6+0.7}_{-2.2-0.5}$
$\frac{\mathcal{B}\left(B^0_s \to K^{*-}\pi^+\right)}{\mathcal{B}(B^0 \to K^{*+}\pi^-)}$	0.9 ± 0.1	0.9 ± 0.3	1.0 ± 0.5
$\mathcal{B}\left(B^0_s \to K^{*+}K^-\right)$	$10.3^{+3.0+4.8}_{-2.2-4.2}$	$6.0^{+1.7+1.7+0.7}_{-1.5-1.2-0.3}$	$8.2^{+1.3+2.1}_{-2.3-2.0}$
$\mathcal{B}\left(B^0_s \to K^{*-}K^+\right)$	$11.3^{+7.0+8.1}_{-3.5-5.1}$	$4.7^{+1.1+2.5+0.0}_{-0.8-1.4-0.0}$	$7.8^{+0.3+1.5}_{-0.5-1.1}$
$\mathcal{B}\left(B^0_s \to K^{*\pm}K^{\mp}\right)$	24 ± 10	12 ± 3	16 ± 3
$\frac{\mathcal{B}\left(B^0_s \to K^{*\pm}K^{\mp}\right)}{\mathcal{B}(B^0 \to K^{*+}\pi^-)}$	2 ± 1	1.4 ± 0.4	1.9 ± 0.4

Table 1.4: Theoretical predictions of relative branching fractions. Values in the average were weighted by the uncertainty on the predictions.

Relative branching fractions	Average value
$\frac{\mathcal{B}\left(B^0 \to K^{*\pm}K^{\mp}\right)}{\mathcal{B}(B^0 \to K^{*+}\pi^-)}$	0.018 ± 0.005
$\frac{\mathcal{B}\left(B^0_s \to K^{*-}\pi^+\right)}{\mathcal{B}(B^0 \to K^{*+}\pi^-)}$	0.90 ± 0.09
$\frac{\mathcal{B}\left(B^0_s \to K^{*\pm}K^{\mp}\right)}{\mathcal{B}(B^0 \to K^{*+}\pi^-)}$	1.7 ± 0.3

2 Experiment

2.1 Introduction

This analysis was performed using data collected during 2011 at the Large Hadron Collider beauty (LHCb) experiment at the Large Hadron Collider (LHC) at the European Organisation for Nuclear Research (CERN). This chapter will describe the LHC and the associated accelerator chain. The LHCb detector, its design constraints and the LHCb software required for the reconstruction and analysis of LHC events will also be described.

2.2 The LHC machine

General design

The LHC is a superconducting hadron accelerator and collider ring [43]. It is located under the French-Swiss border, outside Geneva, and uses the same 26.7 km underground tunnel that was dug for the previous Large Electron Positron (LEP) collider. The tunnel is not fully circular, but is actually composed of eight arcs (for bending) and eight straight sections (for utility or experimental insertions).

Unlike LEP, which was an e^+e^- collider, the LHC uses two proton beams circulating in opposite directions. This requires that the two beams do not share

the same phase space. However the use of specially designed quadrapole magnets allows the beams to share the same beam pipe vacuum. The beams were designed to contain 2808 bunches each, with a nominal bunch spacing of only 25 ns. These two beams are collided at four Interaction Points (IP) around the ring, which are instrumented with four separate experiments: ATLAS [44] at Point 1, ALICE [45] at Point 2, CMS [46] at Point 5 and LHCb [47] at Point 8. The other four points are not instrumented and are used for other insertions such as beam dumps, RF systems and collimators, see Fig. 2.1.

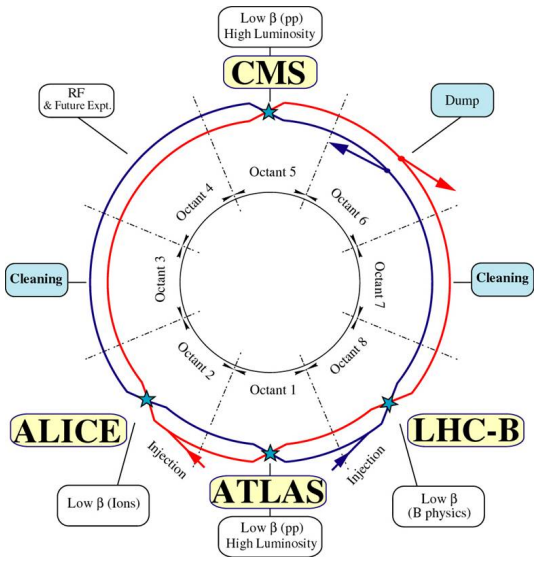

Figure 2.1: Layout of the LHC [48].

The LHC is currently the highest energy particle accelerator in the world, reaching a centre-of-mass energy of 8 TeV in the 2012 running and has a maximum design centre-of-mass energy of 14 TeV. The high energy allows the LHC to potentially create previously inaccessible massive particles, such as the supersymmetric partners of quarks and Gluons, or a high mass Higgs boson.

A high energy is not the only requirement for the physics programme at the LHC. In order to produce large enough samples of different decays for analysis, a high luminosity ($\mathcal{L}$) is required. The LHC was designed to deliver peak luminosities of $\mathcal{L} = 1 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ at the two General Purpose Detectors (GPDs), ATLAS and CMS, and $\mathcal{L} = 2 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ at LHCb. As can be seen in Fig. 2.2, the average instantaneous luminosities achieved by 2011 at LHCb were above the design specifications.

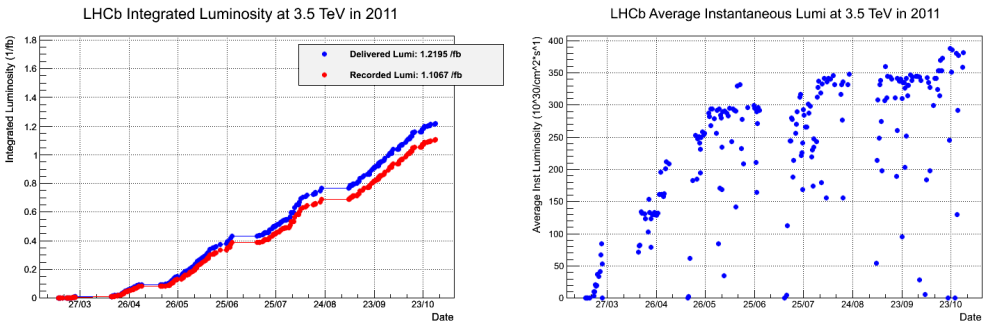

Figure 2.2: Left: LHCb integrated luminosities for 2011. Right: LHCb average instantaneous luminosity during 2011 [49, 50].

Accelerator chain

The LHC ring is not the only stage in the acceleration of the protons, in fact there are multiple different accelerators of increasing energies. The CERN injector chain begins with the stripping of electrons from hydrogen atoms to acquire protons. These protons are injected into the LINear ACcelerator 2 (LINAC2) and accelerated to 50 MeV. From here the beam is moved to the Proton Synchrotron Booster (PSB) where it is split into four separate beams and accelerated to 1.4 GeV. Then the beams are transferred to the Proton Synchrotron (PS) where they are increased to 25 GeV and recombined into two beams. The final stage before the main LHC ring is the Super Proton Synchrotron (SPS) which accelerates the two proton beams to 450 GeV before injecting into the LHC. Fig. 2.3 shows a schematic of the CERN accelerator complex, including both the LHC injector chain and other major experiments.

2.3 The LHCb detector

General design

As mentioned previously, the LHCb detector is located at Point 8 of the LHC ring. The underground cavern used to house the detector was previously used by the DELPHI experiment at LEP. LHCb is a single-arm spectrometer which covers the forward (positive z) region with a pseudorapidity acceptance of $2 \lesssim \eta \lesssim 5$. LHCb uses a right-handed coordinate system with the z axis along the beam and upwards as the y axis. The layout of the LHCb detector is shown in Fig. 2.4, with the interaction region in the VErtex LOcator (VELO) sub-detector.

GPDs at the LHC perform direct searches for new particles produced at the

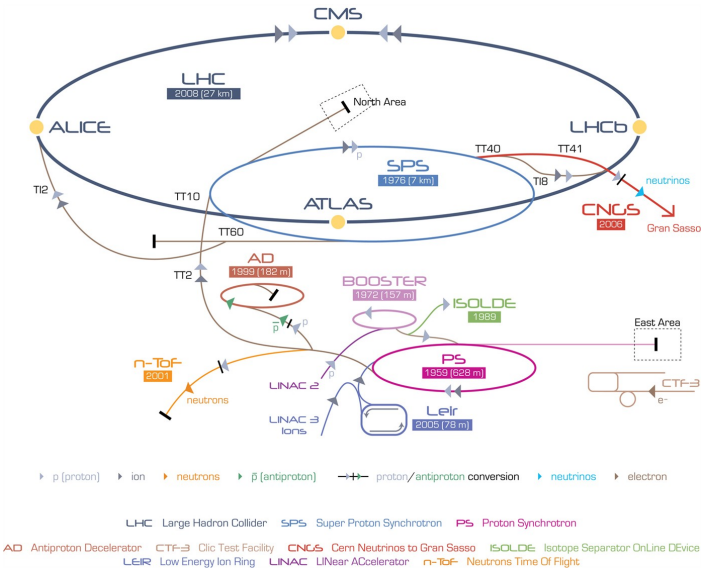

Figure 2.3: Diagram of the LHC complex [51].

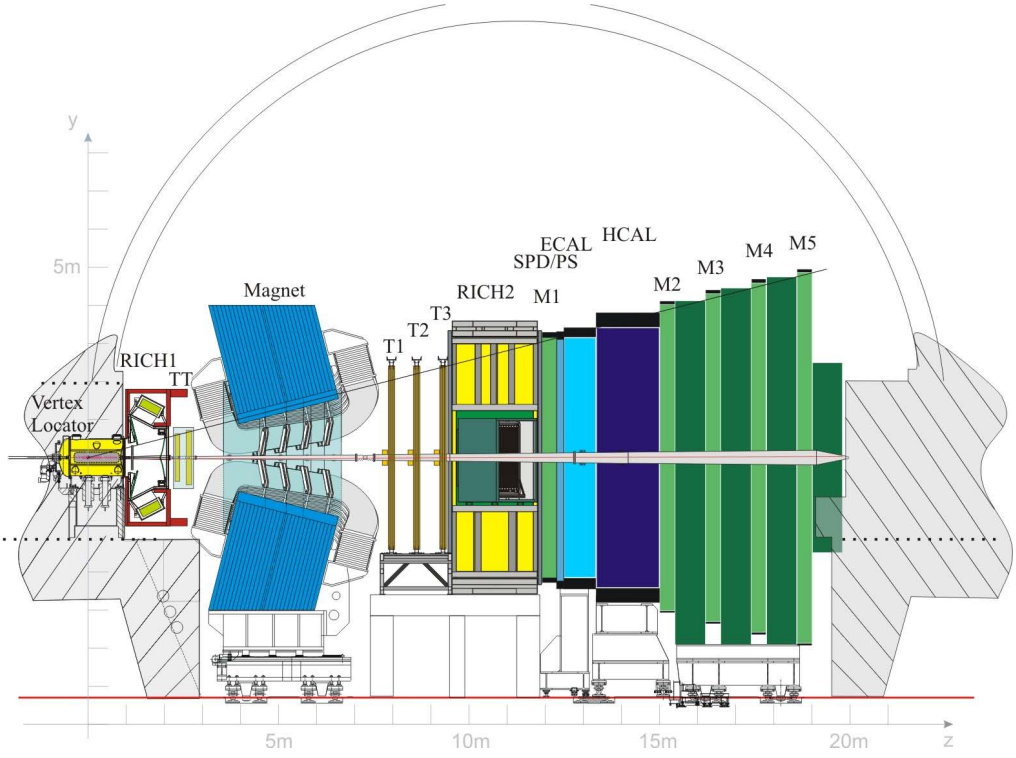

Figure 2.4: Vertical cross section of the LHCb detector [47].

IP due to the large centre-of-mass energies. However LHCb is a precision measurement experiment. It's primary goal is to indirectly search for new physics by making precise tests of the Standard Model predictions of *CP* violation in rare decays of beauty and charm hadrons [47]. Making these precise measurements requires a large dataset of B meson decays. The LHC collisions are the largest source of B mesons in the world, with a $b\bar{b}$ production cross section of $\sim 500\,\mu b$ at the design energy of $\sqrt{s}$ =14 TeV. This $b\bar{b}$ production is predominately in the forward and backward cones around the beam axis. Therefore the design and acceptance of LHCb is specifically tailored to reconstruct the B mesons produced in the forward direction, as shown in Fig. 2.5. Space and cost limitations due to the reuse of the DELPHI cavern prevented the backward region being instrumented.

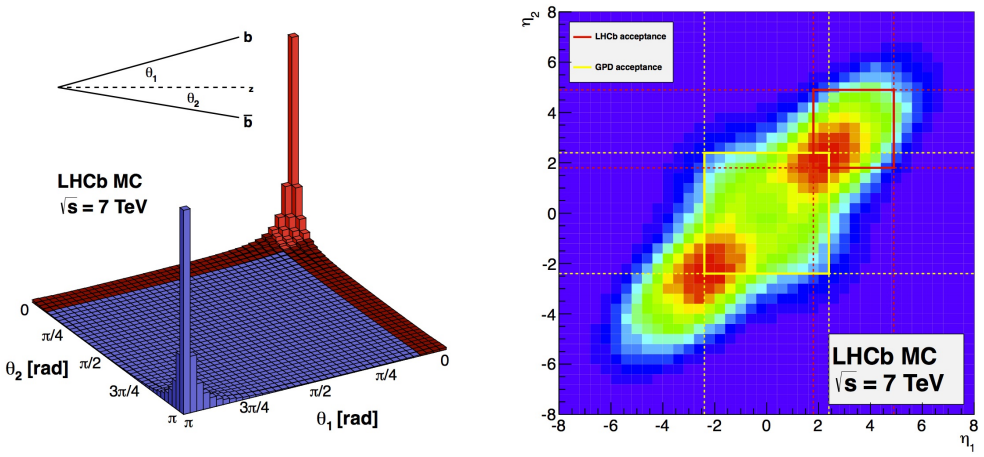

Figure 2.5: Simulated $b\bar{b}$ production angles at LHCb, at $\sqrt{s}$ =7 TeV (produced with PYTHIA 8). Left: The forward and backward $b\bar{b}$ production fractions as a function of the angles with respect to the z-axis, with the LHCb acceptance shaded in red. Right: The $b\bar{b}$ production fractions as a function of the pseudorapidities. The LHCb acceptance is bounded in red and the typical GPD acceptance is bounded in yellow. Red indicates the highest rate of production [52].

Luminosity levelling

One of the major differences between LHCb and the GPDs is its running conditions, specifically the lower design luminosity mentioned in Section 2.2. The LHC has a nominal luminosity of $10^{34}\,cm^{-2}\,s^{-1}$, while LHCb was designed to run at $2 \times 10^{32}\,cm^{-2}\,s^{-1}$. Therefore the levelling technique described in Fig. 2.6 is used to reduce the luminosity. Nevertheless for most of the 2011 data LHCb ran above its

design luminosity, at approximately $3 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. Also Fig. 2.2 shows that in the 2012 the luminosity was closer to $4 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, double the design value!

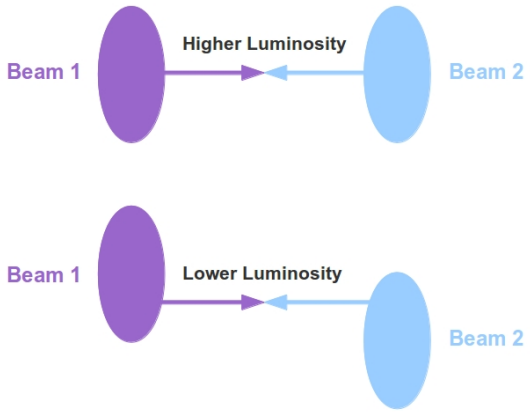

Figure 2.6: A cartoon of the luminosity levelling technique. Top: Beams collide head on, as in the GPDs for maximum luminosity. Bottom: In LHCb the beams are deliberately misaligned to produce a tunable luminosity.

Naively, one might think that reducing the luminosity in this way, would be detrimental to the physics program. With larger luminosities more data could be collected, leading to lower statistical uncertainties. However, there are four main disadvantages to higher luminosities.

Firstly, at higher luminosities more radiation passes through the detector. This can lead to premature breakdown of many components of the detector, leading to reduced performance in a shorter time. This is a large problem for components like the VELO which have to be placed very close to the interaction region.

Secondly, higher luminosities would mean an increased number of proton-proton interactions per bunch crossing. For LHCb, reconstructing the Primary Vertex (PV) where the proton-proton interaction occurred is extremely important for the reconstruction of displaced B meson decay vertices. If the events have many PVs the reconstruction takes longer per event. With multiple PVs it can become ambiguous which PV a B meson originated from and so a 'best' PV must be chosen. During the offline reconstruction, this is calculated as the PV which has the smallest χ^2 of the impact parameter with the B candidate.

Thirdly, at the GPDs the luminosity is not constant during a fill. The luminosity starts at a maximum and slowly reduces as the proton bunches are collided. Due to the luminosity levelling, LHCb receives an almost constant luminosity

throughout a fill, making some calibration and analysis tasks easier at LHCb.

Fourthly, the Level-0 trigger has an output rate limit of $\sim 1.1\,\text{MHz}$. At high luminosities the trigger efficiency on signal events decreases. This means that the final signal event yields do not necessarily scale up with increasing luminosity.

The VELO sub-detector

As mentioned previously, LHCb's main focus is the analysis of charm and beauty hadron decays. Since the $b\bar{b}$ pairs are produced boosted along the beam axis (Fig. 2.5) the B mesons travel $\sim$10 mm before decaying into lighter particles. The VELO is a silicon strip tracker that is designed to provide precise vertex measurement of the PVs and these displaced vertices by recording the tracks close to the IP [53].

To achieve this goal the first active strip of the VELO is only 8.2 mm from the beam during collisions. This is within the LHC's beam aperture during the injection phase of a fill and so the VELO was designed to move the sensors out to 30 mm from the interaction region during the beam setup. Once the beams are declared stable the VELO closes and centres itself around the beam using the Beam Position Monitors. The open and closed positions of the sensor planes are shown in Fig. 2.7. The sensors overlap when the detector is closed. This allows the full ϕ range to be covered and for alignment studies of the halves to be performed using tracks that pass through both halves.

Due to the VELO's close proximity to the beam it is necessary to shield the VELO electronics from the EM field produced by the circulating protons. This is achieved by placing a shaped $\sim 300\,\mu\text{m}$ thick aluminium foil, known as the RF foil, between the sensors and the beam vacuum. Its shape flows around the sensor planes, see Fig. 2.10, and was designed to cause minimal interaction with the tracks prior to them depositing charge in the VELO sensors.

The VELO measures the coordinates of charge deposition using two different sensor planes, one for the radial (R) distance from the beam and one for the azimuthal (ϕ) coordinate around the beam. These two types of sensors are placed back to back in 42 separate modules. The layout of the strips on each type of sensor is sketched in Fig. 2.8.

In order for the VELO to reconstruct tracks covering the entire downstream LHCb acceptance, the particle tracks are required to traverse at least three sensor modules before leaving the VELO. Tracks at the inner and outer limits of the acceptance range should still pass through three modules. As shown in Fig. 2.7, the z spacing of the modules is smallest close to the interaction region so that high p_T

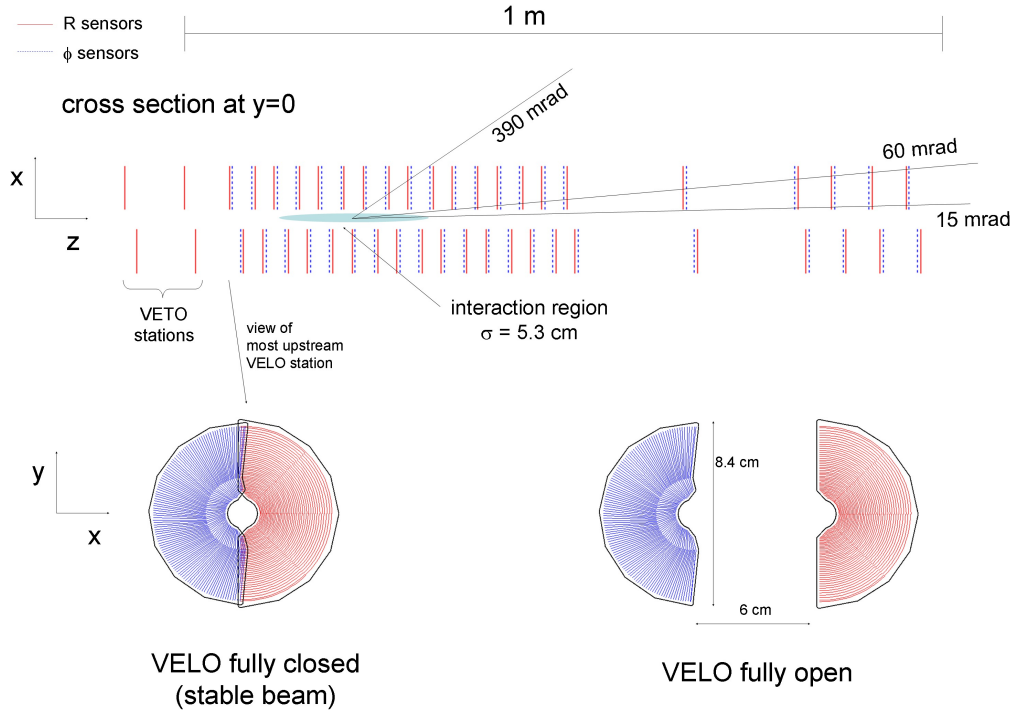

Figure 2.7: VELO (x,z) cross section showing sensor plane z axis positions. Also showing the open and closed VELO positions [54].

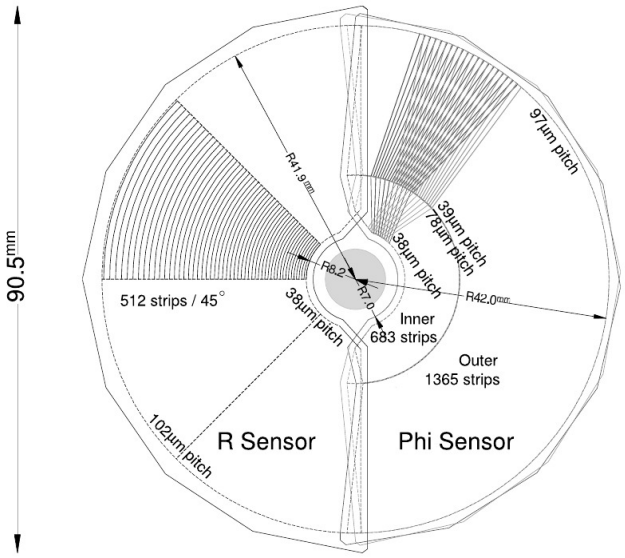

Figure 2.8: Sketch of the VELO R and ϕ sensor planes and their strip orientations (only a fraction of the strips are shown for clarity). The ϕ strips from two adjacent modules are shown to highlight the stereo angle [47].

tracks can still pass through multiple stations. The modules at the furthest downstream end of the VELO are also close together, to maximise the number of modules that very low p_T tracks will traverse. There is larger spacing between the set of upstream and downstream stations as tracks entering this area from the interaction region will have already encountered at least three upstream sensors. Therefore, there is no need to increase the number of sensors in this region.

It is necessary to use a minimal number of sensors in the VELO to reduce the cost, complexity and material budget of the sub-detector. The material budget is a measurement of the amount of material that is traversed by tracks in the sub-detector. Having a high material budget causes the tracks to deposit more energy in the sub-detector, which causes deflection of the tracks and potentially prevents them being measured in the other sub-detectors further downstream.

Fig. 2.9 gives an overview of the material budget distribution in the VELO, measured as the radiation length X_0. Here X_0 is the average distance travelled in a material by an electron of energy E_0 before its energy has reduced to $\frac{E_0}{e}$. The material budget is dominated by the RF foil and sensor planes. Therefore using only a minimal number of sensors is very beneficial to minimising the material budget.

It is possible to use the VELO to image itself by reconstructing material interaction vertices. An early part of the Ph.D studies leading to this thesis involved

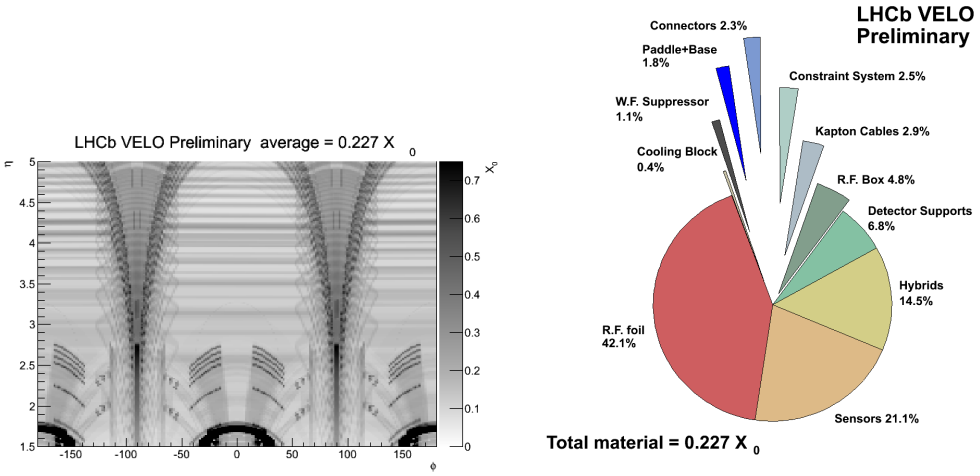

Figure 2.9: Plots describing the material budget of the VELO. Both plots measure the material budget as a function of X_0 from the interaction point to z = 835 mm. Left: Material budget of the VELO as a function of η and ϕ. Right: Percentage of total X_0 for each component of the VELO [55].

the production and analysis of these vertices for a comparison of the material budget of the real VELO and a simulated version. These vertices are shown in Fig. 2.10 for some of the upstream sensors and the surrounding RF foil.

Tracking

The VELO is the first stage of tracking in LHCb. However, there are two more sections which contribute to tracking, the Silicon Tracker (ST) and the Outer Tracker (OT), which is a drift time detector. The ST is itself split into two separate detectors, the Tracker Turicensis (TT) and the Inner Tracker (IT). Fig. 2.11 shows the relative layout of the different ST and OT detectors with all other sub-detectors removed for clarity. The VELO is not in the magnetic field and therefore does not provide any momentum measurement on the tracks, while the ST and OT do.

Both the TT and the IT use silicon micro-strip sensors with a strip pitch of approximately 200 µm. The TT is located upstream of the dipole magnet (Fig. 2.4) and covers the entire LHCb angular acceptance. The IT is made up of three tracking stations inside the T1, T2 and T3 detectors downstream of the magnet. It covers a 120 cm by 40 cm cross shaped region at the centre of each of these three stations (Fig. 2.11). All four of the silicon trackers are consist of four detection layers. The first and last of these layers are arranged vertically while the second and third layers

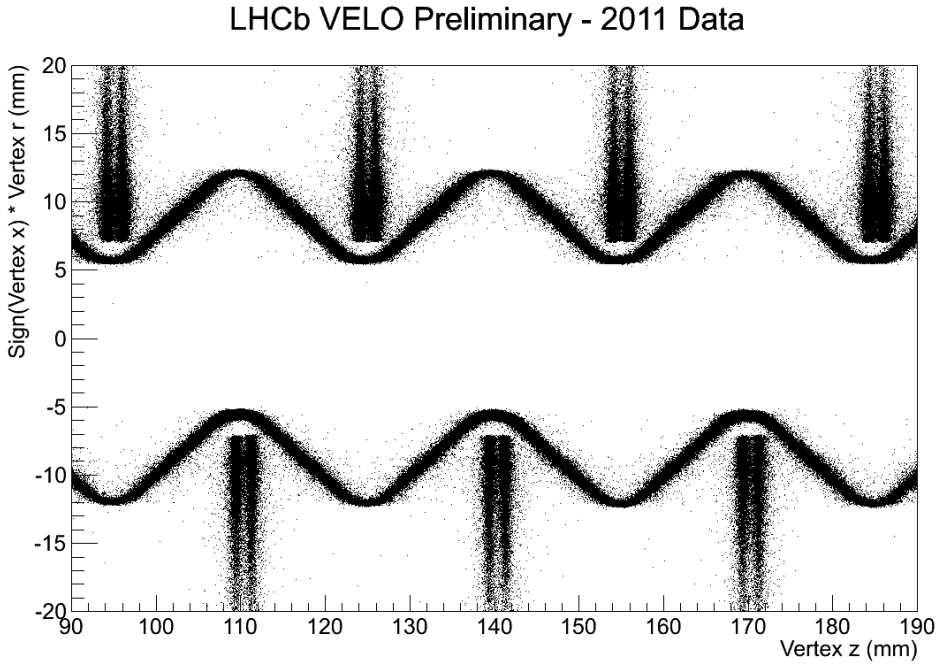

Figure 2.10: Material interaction vertices in the upstream sensors and RF foil of the VELO. Plotted in (z,R) coordinates, where the sign of R is given by the sign of the x component of R [55].

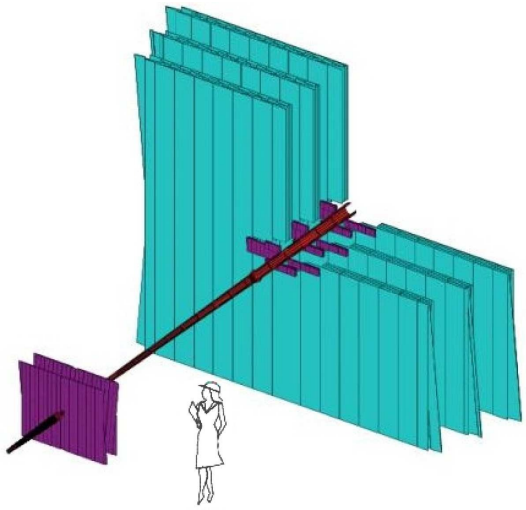

Figure 2.11: The relative layout of the ST (purple) and OT (cyan). The TT is on the left and the IT is on the right, inside the OT [47].

are rotated by stereo angles of $\pm 5°$. This is similar to the VELO's skewed ϕ strips (Fig. 2.8).

The OT was designed to allow both tracking and momentum resolution over a large acceptance area. The OT comprises an array of gas tight straw-tube modules each containing two layers of 4.9 mm drift tubes. It uses a mixture of 70% Ar and 30% CO_2 as its counting gas, to ensure drift times below 50 ns. The three OT stations are each surrounding an IT station (Fig. 2.4).

Ring imaging Cherenkov sub-detectors

In the LHCb detector there are two Ring Imaging CHerenkov (RICH) sub-detectors, RICH1 and RICH2. These are used for particle identification, predominantly the separation of pion and kaon tracks. However, proton, muon, and electron tracks can also be separated with varying degrees of success at different momenta (Fig. 2.12).

RICH1 is located upstream of the dipole magnet (Fig. 2.4) and covers the entire LHCb acceptance. It is used to identify charged particles which are in the momentum range 1-60 GeV. RICH1 uses aerogel and C_4F_{10} radiators to produce the Cherenkov light. To reduce the material budget in the LHCb acceptance both flat and spherical mirrors are used to focus the emitted light onto optical components outside the acceptance.

RICH2 is located downstream of the dipole magnet (Fig. 2.4) and is used to identify higher momentum charged particles of 15 GeV to around 100 GeV. It uses a CF_4 radiator to produce the Cherenkov light and also uses flat and spherical mirrors to reflect this light out of the acceptance. RICH2 has a reduced acceptance of ± 100 mrad vertically and ± 120 mrad horizontally.

In both RICH detectors the Cherenkov light is detected by Hybrid Photon Detectors (HPDs), in the wavelength range 200-600 nm. A typical RICH1 display with reconstructed rings is shown in Fig. 2.12.

Magnet

The magnet used in the LHCb detector is a warm dipole magnet with an integrated field of 4 Tm. It is used to measure the momentum of charged particles from analysis of the amount of bending while particles are inside the field. The magnetic coils are conical saddle shaped and are placed mirror symmetric to one another (Fig. 2.4 and Fig. 2.13). The magnet is water cooled, the coils are made from pure aluminium, and the yoke is made from plates of laminated low carbon steel. The polarity of the magnetic field is reversible, allowing the study of detector asymmetries.

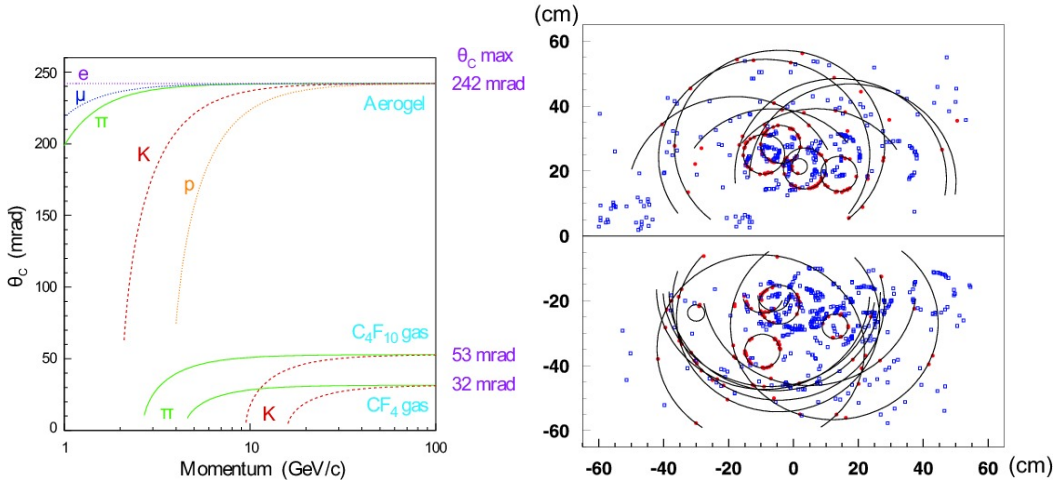

Figure 2.12: Left: The theoretical Cherenkov angle vs. track momentum for different RICH radiators. Right: A typical LHCb event in RICH1 [47].

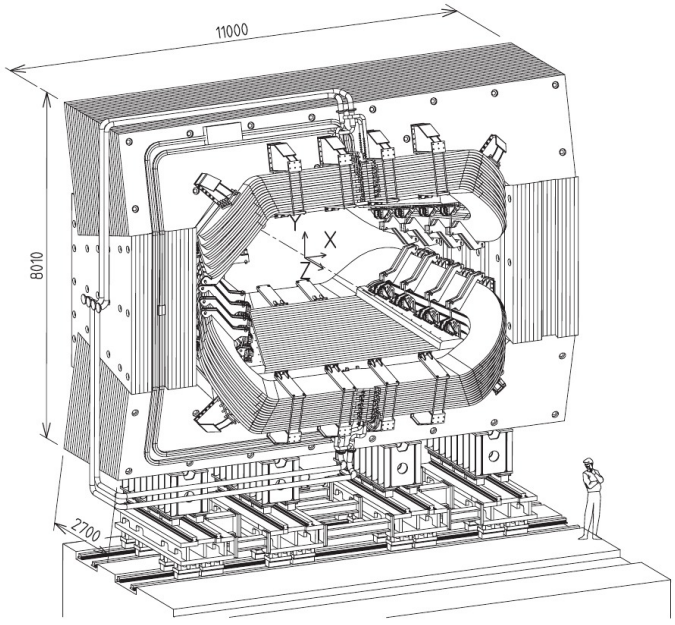

Figure 2.13: Sketch of the LHCb magnet and immediate water cooling connections. The IP is behind the magnet [47].

Calorimeters

The main uses of calorimeters in LHCb are:

1. Identification of high transverse energy clusters for the first level trigger (L0).
2. The reconstruction of electrons, photons and hadrons with measurements of their positions and energies.

There are 4 main calorimeter detectors in LHCb, the Scintillator Pad Detector (SPD), PreShower Detector (PS), the Electromagnetic CALorimeter (ECAL) and the Hadron CALorimeter (HCAL). All of the calorimeters vary the channel density in the (x-y) plane so that the density is higher, closer to the beam (Fig. 2.14). This produces approximately constant angular resolution across the acceptance.

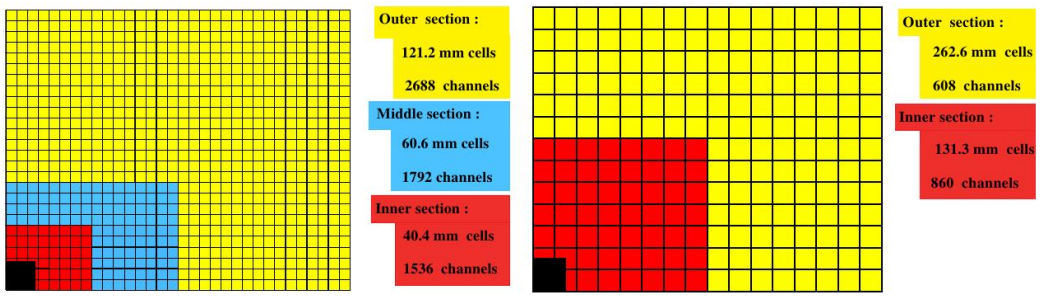

Figure 2.14: Segmentation of one quarter of the LHCb calorimeters and cell dimensions. Black area indicates the space for the beam pipe. Left: SPD, PS, ECAL segmentation with cell dimensions from the ECAL. Right: HCAL segmentation and dimensions [47].

Only charged particles interact in the SPD, allowing photons and electrons to be distinguished by matching deposition in the PS with that in the SPD. The SPD is also used to approximately count the number of charged tracks in the event.[I]

To distinguish between charged hadrons and electrons, the longitudinal electromagnetic shower detection is split into the PS and ECAL. They are placed one after the other (Fig. 2.4) with a lead shield/converter between them and use the shower length/shape to separate the hadrons from electromagnetic particles. The ECAL is required to fully contain the EM showers from high energy photons in order to have excellent energy resolution. It was therefore designed to be 25 radiation lengths thick.

[I]Only approximately due to the magnet sweeping some charged tracks out of the acceptance.

The HCAL's primary function is to identify high transverse momentum hadrons, which is a signature of an event containing a B decay. This information is quickly passed to the L0 trigger so that the event is recorded or discarded. As the trigger does not require the full energy of the hadron, the HCAL does not need to contain the hadronic showers. Therefore to save space it is only 5.6 radiation lengths thick.

The technology of all of these calorimeters is very similar. They use particles exciting a scintillator to produce light, which is then passed to a Photo-Multiplier Tube (PMT) allowing a measurement of the deposited energy.

Muon system

The detection of muons is very important for LHCb as many *CP* sensitive channels have muons in the final state, such as $B^0 \to J/\psi\,(\mu^+\mu^-)K^0_S$. However, high energy muons are very penetrating and thus require thick and multilayered detectors. The muon system consists of five stations (M1-M5) of which four are located downstream of the calorimeters (Fig. 2.15). M2-M5 are interleaved with 80 cm thick iron absorbers to select high energy muons. The minimum energy for a muon to cross all five stations is 6 GeV. M1 is placed before the calorimeters to measure high p_T muons which may not cross all of the calorimeter/muon system before exiting the detector. The p_T information gathered by M1 is also passed to the muon trigger for an improved measurement of the p_T.

The detector technology in all but the inner region of M1 is multi wire proportional chambers (MWPCs) using gas mixtures of $Ar/CO_2/CF_4$. These achieve a time resolution of around 20 ns at voltages of 2.6 kV.

In the inner region of M1, Gas Electron Multipliers (GEMs) are used due to the particle rate being too high for the MWPCs. These are similar in principle to MWPCs as they require the ionization of a gas and multiplication of a charge onto an anode.

Simulation

At LHCb the software package used to create simulated data is called Gauss. In order to generate events, the proton-proton collisions must be simulated first and then the created particles must be decayed. At LHCb, Pythia is used to simulate the collisions and EvtGen is used to decay the particles. To generate a specific decay, EvtGen is configured using a decay file which informs EvtGen of the particles required to be in the decay.

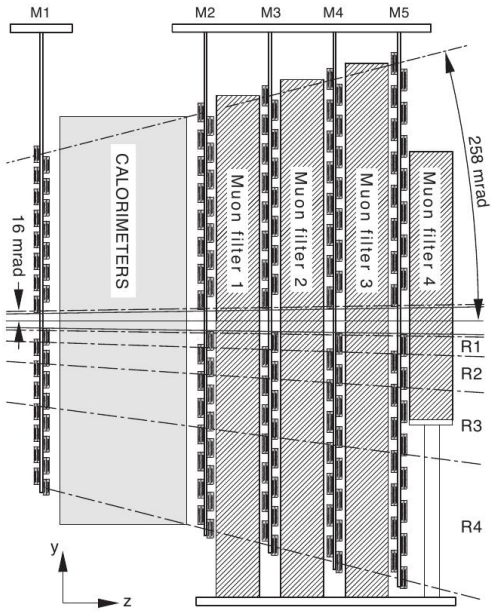

Figure 2.15: Side view of the LHCb muon system [47].

After the particles are generated, they must then be propagated through a simulated version of the LHCb detector, so that the detector response can be simulated. GAUSS uses GEANT4 and a Detector Description DataBase (DDDB) to simulate the particles traversing LHCb. The DDDB is essentially a computer model of LHCb. It stores information about the size, shape and materials that the detector is made of. GEANT4 then propagates the particles through this geometry.

The package BOOLE is the final stage in the LHCb simulation. It takes output of the previous stages and simulates the detector response, meaning that it simulates the digital output of the real detector.

Trigger

Not every collision at the LHCb IP will contain a decay that is of interest to the physics program. During the 2011 running, the rate of interactions at LHCb was $\sim 10\,\text{MHz}$ (rather than the 40 MHz design value). If LHCb had to fully analyse and store every event, the cost in computing resources and time would be be prohibitive. In order to be able to store the events, the rate must be reduced to $\sim 3\,\text{kHz}$. Therefore LHCb employs a trigger system which can reject uninteresting events quickly, using different levels of reconstruction and complexity.

The LHCb trigger is split into two major levels, the Level-0 (L0) hardware trigger is first and the High Level Trigger (HLT) implemented as software is second. A simple flow chart of the rate and type of accepted events moving though the trigger system is shown in Fig. 2.16.

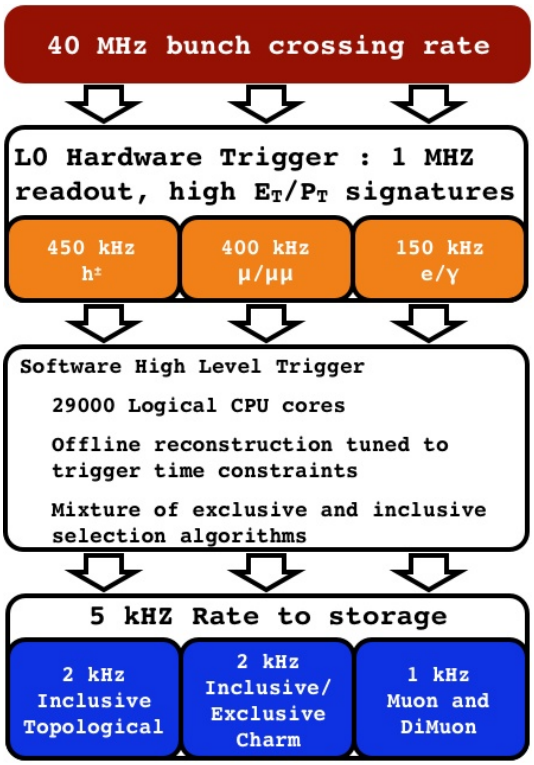

Figure 2.16: Flow chart of the LHCb trigger system [52].

The L0 hardware trigger collects information from the hardware calorimeter trigger and hardware muon trigger to make its decisions. High E_T electrons, photons or hadrons are selected by the calorimeter trigger, while high p_T muons are selected by the muon trigger. Events pass through the L0 trigger at about 1 MHz.

The HLT is split into two parts, HLT1 and HLT2. The HLT1 uses the tracking information from the VELO, ST, OT and muon systems to reduce the passing rate down to ~ 30 kHz. It does this by selecting b-hadron candidates which have large impact parameters with respect to the Primary Vertex (PV) and which have high p_T. The HLT1 can also apply Global Event Cuts (GECs) on variables such as the number of VELO hits or SPD multiplicity. The SPD multiplicity gives an approximate measurement of the number of charged tracks in the event. If this is large the reconstruction of the event is time consuming, less efficient and will have

larger uncertainties in general. Placing reasonable requirements on the maximum multiplicity is an excellent way to decrease the average processing time without sacrificing useful events.

The HLT2 is the final stage of the trigger and uses a simple event reconstruction to apply more complex selection requirements. Simple track combinations are created for particles such as ϕ. These are then used to make specified final states, allowing loose requirements on the B mass and momentum direction. The HLT2 then writes out to disk for offline analysis at a rate $\lesssim 5\,\text{kHz}$.

Using software triggers at LHCb rather than hardware triggers has several benefits. As long as the calculations are fast enough, the software can have access to the whole event. This allows the trigger to make more complex and varied selections more easily than a hardware one could. Software triggers can also be changed much more easily than hardware ones if the physics requirements or conditions of the experiment change.

When performing an analysis, the software trigger chain can be re-applied by running the MOORE package. It is also this package which runs the trigger on the simulated events from BOOLE.

Reconstruction and analysis

To do the full event reconstruction, LHCb uses the package BRUNEL. It can take either the output from BOOLE or the LHCb Data AcQuisition (DAQ). BRUNEL fits tracks and finds vertices allowing the calculation of particle properties.

DAVINCI is the LHCb physics analysis package. Tracks and vertices from BRUNEL are combined and selections based on the decay chains of interest can be applied. It applies the Decay Tree Fitter method, previously used at BABAR, to recursively fit the decay tree and apply mass constraints [56]. The properties of the particles satisfying specified requirements in each event can be written out to a ROOT ntuple.

Stripping lines

The full number of events coming from the recorded data at LHCb is enormous and the vast majority of them contain no decays of interest for any single analysis in the LHCb physics programme. This makes it computationally expensive for every analysis to needlessly construct common particle hypotheses and reject the events separately. Therefore, the LHCb experiment centralises this process by allowing the many different analyses to submit the selection requirements, known as *Stripping*

Lines.

During the central stripping process, after an event is reconstructed it can be checked against all of the stripping lines in a much more efficient manner. This allows only the events passing one or more of the lines to be retained, and each event is flagged as to which line or lines it has passed. The reduced dataset contains only decays of interest to the physics programme, flagged by stripping line and reconstructed. Additionally, stripping lines for similar physics programmes are grouped into *Streams* and the final datasets are split according to these streams. This vastly decreases the amount of time required to extract the information about the signal decays for the end user. There is also the benefit of different analyses using the same datasets, allowing more direct comparison of results.

3 Selection

3.1 Introduction

A fundamental part of the analysis is the selection of events containing the signal decays. A selection is in general designed to maximise the number of retained signal events, while rejecting events not containing signal decays.

In Section 3.2 the data sample used in the analysis will be described. Then in Sections 3.3-3.4 the initial selection requirements applied during and immediately after the reconstruction of the decays will be listed. The specific vetoes applied to remove the charm backgrounds are given in Section 3.5. A multivariate analysis was used, the method by which this was trained and optimised is detailed in Section 3.6. As measurement of the $B^0_{d,s} \to K^{*\pm}h^{\mp}$ decays requires the identification and separation of pion and kaon tracks, particle identification requirements were optimised and applied. These are described in Section 3.7. The fits to the data samples, detailed in Chapter 5, require that the data be fitted to observable parameters within specified ranges. These parameters and the ranges used are described in Section 3.8.

3.2 Data sample

This analysis uses a data sample with an integrated luminosity of $\int \mathcal{L} = 1025 \pm 36$ pb^{-1}. The data was recorded during 2011 operation of the LHCb experiment. The recorded integrated luminosity for the 2011 period can be seen in Fig. 2.2.

3.3 Initial selection

The $B^0_{d,s} \to K^0_S h^\pm h'^\mp$ stripping line

Prior to this analysis, a measurement of the branching fractions of the six $B^0_{d,s} \to K^0_S h^\pm h'^\mp$ modes was made at LHCb [1]. This required the creation and optimisation of an inclusive stripping line, `B2KShh`, to select all six modes.

As there are finite computing resources available for both processing and storage of stripped events, there were competing requirements placed on the line. It would be preferable to place only very loose requirements on the decay properties, to maximise the retained signal. However, the central stripping required that the retention rate be $\lesssim 5 \times 10^{-4}$event^{-1}, and an event processing speed of $\lesssim 1$ ms per event.

In order to meet these requirements, cuts on certain variables had to be made. The choice of the cut values and which variables to cut on was performed by taking the $B^0 \to K^0_S \pi^+ \pi^-$ simulated events and a sample of upper sideband collision events. Then the variables were plotted in both samples allowing the most discriminating variables to be found. Then the cuts were initially chosen based on those from previous 3-body analyses at LHCb and refined by eye. The $B^0_{d,s} \to K^0_S h^\pm h'^\mp$ simulated events were then used to check that the final cuts did not cause extremely low efficiency in the physically interesting edges and corners of the Dalitz plot. The variables and their requirements are discussed below.

The fastest cuts to process are GECs, which only assume minimal track reconstruction and are usually measures of the particle density of the event. One such variable is number of `Long` tracks, which are tracks that have hits in both the VELO and in the tracking sub-detectors downstream of the magnet. In the `B2KShh` lines, it is required that the number of `Long` tracks per event is less than 250.

The next fastest cuts to process are placed on variables associated with the reconstructed tracks from the daughter particles. Such as momentum, track χ^2 and the impact parameter with respect to the PV. These are fast because these variables do not require the combination of the daughter tracks into a vertex and the associated calculations. However, the `B2KShh` lines do not apply tight constraints on

the daughter momenta. Doing so would constrain the phase space of the Dalitz plots, preventing analysis of certain regions and resonances near the edges and corners, see Appendix A.

Daughter track requirements

During reconstruction of the event, collections of particle hypotheses from tracks are made for common particle types, such as kaons and pions. In the `B2KShh` lines, the B candidate decay vertex is formed by combining a K^0_S candidate with two oppositely charged tracks. As a track does not have a defined particle ID prior to the PID information being used, at LHCb all tracks are assumed to be pions at first. The charged tracks for the vertex combination come from either the `StdLoosePions` or `StdLooseKaons` collections, depending on the final state. The `StdLoosePions` collection has minimal transverse momentum cuts applied and no PID cuts. While the `StdLooseKaons` is a subset of `StdLoosePions` when a PID cut of $\mathrm{DLL}_{\mathrm{K}\pi} > -5$ is applied, to preferentially select kaons. The total requirements made on the charged pion and kaon tracks, during and after the stripping process, are detailed in Table 3.1.

Table 3.1: Total selection requirements for the charged pion and kaon tracks.

Variable definition	Pion requirement	Kaon requirement
Track fit χ^2 per degree of freedom	< 4	< 4
"Clone distance"[I]	> 5000	> 5000
Kaon vs. pion PID ($\mathrm{DLL}_{\mathrm{K}\pi}$)	-	> -5
Total momentum	$< 100\,\mathrm{GeV}/c$	$< 100\,\mathrm{GeV}/c$
Transverse momentum	$> 250\,\mathrm{MeV}/c$	$> 250\,\mathrm{MeV}/c$
Minimum impact parameter χ^2 *w.r.t.* PVs	> 4	> 4

The requirement that the track momentum is less than $100\,\mathrm{GeV}/c$ was due to the limitations of the RICH detectors. As shown in Fig. 2.12, the separation of the Cherenkov opening angles in the RICH sub-detectors is negligible for track momenta greater than $100\,\mathrm{GeV}/c$. This means that PID requirements are ineffective on these tracks and consequently they are removed.

K^0_S candidate requirements

As the K^0_S has no charge, it doesn't create tracks in the detector. The decay vertex of a K^0_S candidate to $\pi^+\pi^-$ is inferred by the combination of two pion tracks, allowing an invariant mass and momentum calculation for the K^0_S candidate. The K^0_S candidates are split into two collections, Long-Long K^0_S (`StdLooseKsLL`) and Down-Down K^0_S (`StdLooseKsDD`). The Long-Long K^0_S are formed from pion candidates in the `StdLoosePions` collection mentioned above, these only use `Long` pion tracks. While the Down-Down K^0_S use `StdNoPIDsDownPions`, which are tracks that have no hits in the VELO because the K^0_S decayed further in the z direction than the VELO extends. This separation is necessary because the lack of VELO tracks leads to very different mass resolutions, lifetimes and reconstruction effects between the K^0_S candidates. The samples therefore have different distributions in many variables, potentially requiring different selections. For example, the expected resolution for the Long-Long K^0_S is $\sim 10\,\text{MeV}/c^2$ while the Down-Down K^0_S candidates have a larger resolution of $\sim 20\,\text{MeV}/c^2$. This makes it possible to put tighter requirements on the mass of the Long-Long K^0_S candidates. The total requirements made on the Long-Long and Down-Down K^0_S candidates are described in Table 3.2.

Table 3.2: Selection requirements for the Long-Long and Down-Down K^0_S candidates.

Variable definition	Long-Long requirement	Down-Down requirement
K^0_S daughter track momentum	$> 2\,\text{GeV}/c$	$> 2\,\text{GeV}/c$
K^0_S daughter track fit χ^2 per degree of freedom	< 4	-
K^0_S daughter minimum impact parameter χ^2 *w.r.t.* PVs	> 9	> 4
Mass difference *w.r.t.* nominal K^0_S mass	$< 20\,\text{MeV}/c^2$	$< 30\,\text{MeV}/c^2$
χ^2 of K^0_S vertex fit	< 12	< 12
χ^2 separation of K^0_S vertex and associated PV	> 80	> 50
K^0_S momentum	-	$> 6\,\text{GeV}/c$

B candidate requirements

During the stripping process a four-momentum combination of the daughter tracks is made. This allows some loose requirements prior to the full vertex fit, which are

[1]The clone distance is calculated as the Kullback-Liebler distance of the track [57].

shown in Table 3.3.

Table 3.3: Selection requirements for the B candidates after four-momentum combination.

Variable Definition	Selection requirement
Transverse momentum of the B candidate	$> 1000\,\mathrm{MeV}/c$
Sum of the daughters' transverse momenta	$> 3000\,\mathrm{MeV}/c$
p_T of at least two B daughters	$> 800\,\mathrm{MeV}/c$
Mass of the B candidate	$4779 < m_{\mathrm{K_S^0 h^+ h^-}} < 5866\,\mathrm{MeV}/c^2$
IP *w.r.t.* PV of highest p_T B daughter	$> 0.05\,\mathrm{mm}$
Maximum DOCA χ^2 of any 2 daughters	< 5

After the fit for the B decay vertex more cuts are applied, which are listed in Table 3.4. In these cuts, the pointing angle is defined as the angle between the momentum of the B and the direction vector between the PV and the B vertex. Therefore by requiring that the cosine of this value be close to one, the B particle is required to have come from the associated PV.

The separation of the $\mathrm{K_S^0}$ and B vertices is also required to be at least 10 mm, with the $\mathrm{K_S^0}$ having travelled in the positive z direction. This is only relevant for the Long-Long $\mathrm{K_S^0}$ candidates as the Down-Down $\mathrm{K_S^0}$ ones, by definition, are separated from the B decay by the LHCb magnet. It is necessary to enforce the separation to prevent falsely reconstructed $\mathrm{K_S^0}$. As $\mathrm{K_S^0}$ candidates that come from the negative z direction compared to the B vertex do not make sense due to the boosting of the B production, see Fig. 2.5, but could otherwise pass the selection cuts.

Table 3.4: Selection requirements for the B candidates after the full vertex fit.

Variable Definition	Selection requirement
Transverse momentum of the B candidate	$> 1500\,\mathrm{MeV}/c$
χ^2 of B vertex fit	< 12
Cosine of B pointing angle	> 0.9999
Minimum B IP χ^2 *w.r.t.* PVs	< 8
Minimum vertex distance *w.r.t.* PVs	$> 1\,\mathrm{mm}$
χ^2 separation of B vertex and associated PV	> 50
Separation between the $\mathrm{K_S^0}$ and B vertices in the positive z direction	$> 10\,\mathrm{mm}$

In later versions of the stripping line, some of the requirements were tightened after further analysis showed an increase in discrimination for the Down-Down $\mathrm{K_S^0}$

signal. These were applied to the data sample for this analysis and are listed in Table 3.5.

Table 3.5: Extra selection requirements for candidate decays containing Down-Down K^0_S, added from later stripping versions.

Variable Definition	Selection requirement
Sum of the daughters' transverse momenta	$> 4200\,\text{MeV}/c$
Sum of the daughters' minimum IP χ^2 *w.r.t.* PVs	> 50
Minimum B IP χ^2 *w.r.t.* PVs	< 6

3.4 Trigger requirements

As described in Section 2.3, the LHCb trigger is made up of three major sections, one hardware and two software. The trigger channels applicable to this analysis are the hadronic ones, whose many requirements and performances are more fully described in Ref [58, 59].

Although events are necessarily selected by the trigger before being written to tape, it is also possible to select events after reconstruction of the decay, based on which particle in the event caused the trigger to fire. In this way, it is possible to require that the trigger fired on a particle used in the reconstruction of the selected decay. This is known as a Trigger On Signal (TOS) requirement. If no particle used in the signal decay reconstruction caused the trigger to fire, it is known as a Trigger Independent of Signal (TIS) requirement. The TIS and TOS requirements made on the data sample are listed in Table 3.6 and are the same as those optimised for the $B^0_{d,s} \to K^0_S h^\pm h'^\mp$ analysis.

The L0 trigger cuts in Table 3.6 are used to accept events with heavy flavour decays that decay to hadrons. These trigger cuts require that a hadron with a transverse energy, E_T, greater than 3.5 GeV was recorded in the event. They also remove events with large occupancies in the SPD. The HLT1 trigger cut accepts events where the B decays a significant distance from the associated PV, by applying cuts on the track quality from the VELO reconstruction [59]. Also of note are the HLT2 two, three and four-body topological triggers. There are two types of topological trigger, one called 'Simple' and one which uses a so-called Bonsai Boosted Decision Tree (BBDT) [60].

Table 3.6: Trigger selection requirements after reconstruction of decays.

Trigger Level	Selection requirement
L0	`L0Hadron TOS OR L0Global TIS`
HLT1	`Hlt1TrackAllL0 TOS`
HLT2	`(2,3,4)-BBDT TOS OR (2,3,4)-Simple TOS`

3.5 Charm vetoes

As this analysis is focused on the charmless decays of B mesons, peaking backgrounds containing charmonium particles and D mesons should be excluded where possible. In this section a simple summary of the charm vetoes applied to the dataset is given. For a more thorough treatment of the backgrounds being vetoed see Chapter 4.

Charmonium particles such as J/ψ and $\psi(2S)$ can decay to $\mu^+\mu^-$. To remove these muonic tracks, it is required that the charged tracks from both the B decay vertex and K^0_S decay satisfy the `isMuon==0` requirement. This is a μ-ID variable which is true if the tracks are matched with hits in the muon system.

It is also possible for the different charmonium particles to decay to $\pi^+\pi^-$ or K^+K^- final states. These events are removed by making requirements on the difference between the reconstructed mass of the two charged tracks under the relevant particle hypotheses and the PDG mass of the charmonium particle, as shown in Table 3.7. This procedure is also applied to veto the $B \to Dh$ and $\Lambda_b \to \Lambda_c h$ decays. The specific modes considered are described in Table 3.7. Here the inclusion of the charge conjugate processes is implied.
Studies performed for the $B^0_{d,s} \to K^0_S h^\pm h'^\mp$ analysis showed that these cuts were sufficient to remove all but a negligible amount of these background decays.

3.6 Neural net

Multivariate selections were used to discriminate the $B^0_s \to K^{*\pm}K^\mp$ signal against the combinatorial background. The neural net training software NEUROBAYES was used to train the neural nets and use them to select the signal events. Two neural nets were trained independently for the Long-Long K^0_S and Down-Down K^0_S using simulated events for the signal and $K^0_S\pi^+\pi^-$ upper side-band events for the background. To validate the use of simulated events rather than collision data, the ${}_s\mathcal{P}lots$ method was used.

Table 3.7: Charm vetoes.

Charmed particle and decay mode	Selection requirement
$J/\psi \to \pi^-\pi^+$	$\lvert m_{\rm Rec} - m_{\rm PDG}\rvert > 30\,{\rm MeV}/c^2$
$J/\psi \to K^-K^+$	$\lvert m_{\rm Rec} - m_{\rm PDG}\rvert > 30\,{\rm MeV}/c^2$
$\chi_c \to \pi^-\pi^+$	$\lvert m_{\rm Rec} - m_{\rm PDG}\rvert > 30\,{\rm MeV}/c^2$
$\chi_c \to K^-K^+$	$\lvert m_{\rm Rec} - m_{\rm PDG}\rvert > 30\,{\rm MeV}/c^2$
$D^0 \to K^-\pi^+$	$\lvert m_{\rm Rec} - m_{\rm PDG}\rvert > 30\,{\rm MeV}/c^2$
$D^0 \to \pi^+\pi^-$	$\lvert m_{\rm Rec} - m_{\rm PDG}\rvert > 30\,{\rm MeV}/c^2$
$D^0 \to K^+K^-$	$\lvert m_{\rm Rec} - m_{\rm PDG}\rvert > 30\,{\rm MeV}/c^2$
$D^+ \to K^0_S K^+$	$\lvert m_{\rm Rec} - m_{\rm PDG}\rvert > 30\,{\rm MeV}/c^2$
$D^+ \to K^0_S \pi^+$	$\lvert m_{\rm Rec} - m_{\rm PDG}\rvert > 30\,{\rm MeV}/c^2$
$D_s^+ \to K^0_S K^+$	$\lvert m_{\rm Rec} - m_{\rm PDG}\rvert > 30\,{\rm MeV}/c^2$
$D_s^+ \to K^0_S \pi^+$	$\lvert m_{\rm Rec} - m_{\rm PDG}\rvert > 30\,{\rm MeV}/c^2$
$\Lambda_c \to pK^0_S$	$\lvert m_{\rm Rec} - m_{\rm PDG}\rvert > 30\,{\rm MeV}/c^2$

${}_s\mathcal{P}lots$ method

If a data sample contains both signal and background events, a fit to a discriminating observable can extract the signal and background yields. The ${}_s\mathcal{P}lots$ technique uses these yields, the Probability Density Functions (PDFs) and the correlation matrix of the fit to calculate an ${}_s\mathcal{W}eight$ for each species on a per event basis. These ${}_s\mathcal{W}eights$ are normalised, such that when summed over all events in the sample they give the fitted species yield. Also, for each event the sum of the different species ${}_s\mathcal{W}eights$ is unity.

The calculated signal ${}_s\mathcal{W}eights$ can be used when plotting the distribution of variables not used in the likelihood fit. The signal distribution of that variable is then recovered, so long as the variable is uncorrelated with the fitting observable. This technique is used in this analysis to check that the distributions of discriminating variables in the neural net for simulated and collision events are the same. The mathematics behind the ${}_s\mathcal{P}lots$ technique and its limitations are more thoroughly described in [61].

Training of neural nets

The two main decisions for the training of of a neural net are:

1. Which variables to use in the input layer.

2. Which data samples to use for the signal and background events.

The input variables chosen for the neural nets are shown in Tables 3.9-3.10. These were selected based on previous studies for the $B^0_{d,s} \to K^0_S h^\pm h'^\mp$ analysis showing that they have some discriminating power when compared to the combinatorial background. They are deliberately chosen to be mainly topological variables to reduce Dalitz plot bias and are extensively used in the $B^0_{d,s} \to K^0_S h^\pm h'^\mp$ stripping line.

For the training of a neural net it is necessary to use data samples containing known signal and background events. To avoid potential bias in the neural net, it was decided to train on datasets independent to that containing the $B^0_s \to K^{*\pm} K^\mp$ events. The largest samples of signal and combinatorial events for $B^0_{d,s} \to K^0_S h^\pm h'^\mp$ decays, independent of the $K^0_S K^\pm \pi^\mp$ final state, come from the $B^0 \to K^0_S \pi^+ \pi^-$ simulation and the $K^0_S \pi^+ \pi^-$ collision dataset respectively.

However, there is no guarantee that the distributions of the various input variables in the $K^0_S \pi^+ \pi^-$ signal data will be correctly modelled by the $B^0 \to K^0_S \pi^+ \pi^-$ simulation. Furthermore, it must be verified that the $B^0_s \to K^0_S K^\pm \pi^\mp$ signal distributions are sufficiently similar to the $B^0 \to K^0_S \pi^+ \pi^-$ decays. To check that it is valid to use the $B^0 \to K^0_S \pi^+ \pi^-$ simulation as a replacement signal data sample for the $B^0_s \to K^0_S K^\pm \pi^\mp$ data, it is first necessary to extract a sample of $B^0 \to K^0_S \pi^+ \pi^-$ events from the collision data.

To do this, a simple extended fit model for the signal B invariant mass peak and total background was created. The signal peak was modelled by the sum of two Gaussian PDFs, which share the mean parameters, while the total background was modelled by an exponential PDF with a negative decay constant. All parameters were allowed to freely float in this fit.

In order to show that the signal PDF can correctly fit to the signal peak in the B invariant mass, the model was applied to both the $B^0 \to K^0_S \pi^+ \pi^-$ and $B^0_s \to K^0_S K^\pm \pi^\mp$ truth matched simulation as shown in Fig. 3.1. The fit was then applied to the $K^0_S \pi^+ \pi^-$ data, also shown in Fig. 3.1. All previously described requirements were applied to these data samples prior to fitting, as well as some sensible PID requirements on the charged tracks. The PID requirements were taken from the previous $B^0_{d,s} \to K^0_S h^\pm h'^\mp$ analysis [1] in order to minimise the misidentification background before fitting. The signal yields found from the relevant fits are listed in Table 3.8.

By fixing all of the fit parameters, apart from the signal and background yields, ${}_s\mathcal{W}eight$ values were extracted using the ${}_s\mathcal{P}lots$ method. Then by applying the ${}_s\mathcal{W}eights$ to the $K^0_S \pi^+ \pi^-$ data, the $B^0 \to K^0_S \pi^+ \pi^-$ signal distributions for potential input variables to the neural net were plotted. These histograms were then

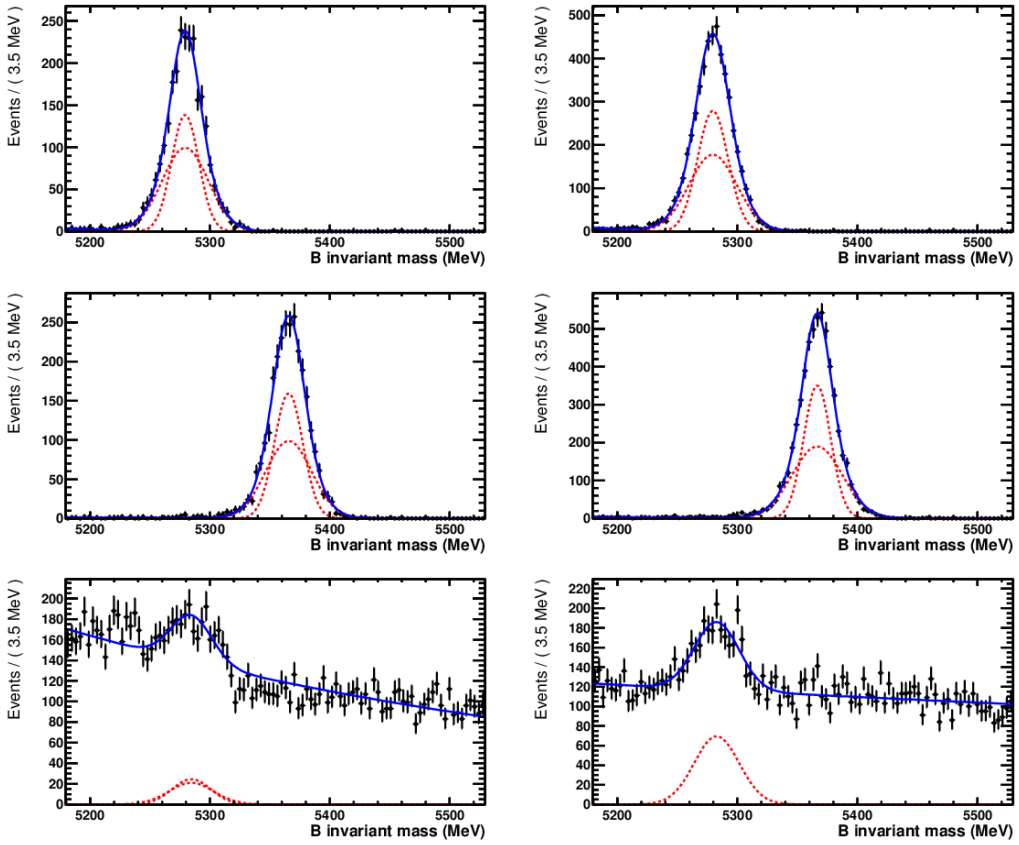

Figure 3.1: Simulation and collision data for $B^0 \to K^0_S\pi^+\pi^-$ MC (top), $B^0_s \to K^0_S K^\pm\pi^\mp$ MC (middle) and $K^0_S\pi^+\pi^-$ collision data (bottom) fitted. Signal modelled with a double Gaussian PDF, total background modelled with an exponential PDF. Left: Long-Long K^0_S. Right: Down-Down K^0_S.

Table 3.8: Signal yields for $B^0 \to K^0_S\pi^+\pi^-$ MC, $B^0_s \to K^0_S K^\pm\pi^\mp$ MC and $B^0 \to K^0_S\pi^+\pi^-$ Collision data

Dataset	Fitted signal yield
Down-Down K^0_S	
$B^0 \to K^0_S\pi^+\pi^-$ MC	5076 ± 73
$B^0_s \to K^0_S K^\pm\pi^\mp$ MC	5658 ± 76
$B^0 \to K^0_S\pi^+\pi^-$ Collision	930.7 ± 2.8
Long-Long K^0_S	
$B^0 \to K^0_S\pi^+\pi^-$ MC	2496 ± 51
$B^0_s \to K^0_S K^\pm\pi^\mp$ MC	2708 ± 52
$B^0 \to K^0_S\pi^+\pi^-$ Collision	561 ± 73

plotted for the $B^0 \to K^0_S \pi^+ \pi^-$ and $B^0_s \to K^0_S K^\pm \pi^\mp$ simulated data. Then χ^2 tests were used to calculate p-values between the data samples for the different input variables. The p-value in this case is the probability that a difference at least as extreme as the one found between the two samples would be produced, even if the underlying physical distributions were the same. Therefore, a low p-value indicates that it is unlikely that the variables have the same underlying distribution between the two samples. The resulting p-values are detailed in Tables 3.9-3.10 and the ${}_s\mathcal{P}lots$ of the variables for the three data samples are shown in Figs. 3.2-3.10.

Table 3.9: p-values of simulation and Data comparisons for neural net variables for Down-Down K^0_S. Comparing $B^0_s \to K^0_S K^\pm \pi^\mp$ and $B^0 \to K^0_S \pi^+ \pi^-$ simulated events, and $B^0 \to K^0_S \pi^+ \pi^-$ simulated and collision events.

Variable Name	MC vs. MC p-value	MC vs. Data p-value
B impact parameter χ^2	0.61	0.065
B end vertex χ^2	0.42	0.61
B transverse momentum	0.48	0.14
$\log_{10}$(1 - cosine of B pointing angle)	0.18	0.88
$\log_{10}$(B vertex distance χ^2)	0.26	0.27
$\log_{10}$(K^0_S IP χ^2)	0.86	0.61
$\log_{10}$(K^0_S end vertex χ^2)	0.31	0.41
$\log_{10}$(negative track IP χ^2)	0.30	0.42
$\log_{10}$(positive track IP χ^2)	0.20	0.79

Table 3.10: p-values of simulation and Data comparisons for neural net variables for Long-Long K^0_S. Comparing $B^0_s \to K^0_S K^\pm \pi^\mp$ and $B^0 \to K^0_S \pi^+ \pi^-$ simulated events, and $B^0 \to K^0_S \pi^+ \pi^-$ simulated and collision events.

Variable Name	MC vs. MC p-value	MC vs. Data p-value
B impact parameter χ^2	0.57	0.04
B end vertex χ^2	0.34	0.15
B transverse momentum	0.51	0.12
$\log_{10}$(1 - cosine of B pointing angle)	0.26	0.0049
$\log_{10}$(B vertex distance χ^2)	0.34	0.067
$\log_{10}$(K^0_S IP χ^2)	0.62	0.26
$\log_{10}$(K^0_S end vertex χ^2)	0.25	0.81
$\log_{10}$(negative track IP χ^2)	0.59	0.27
$\log_{10}$(positive track IP χ^2)	0.053	0.67

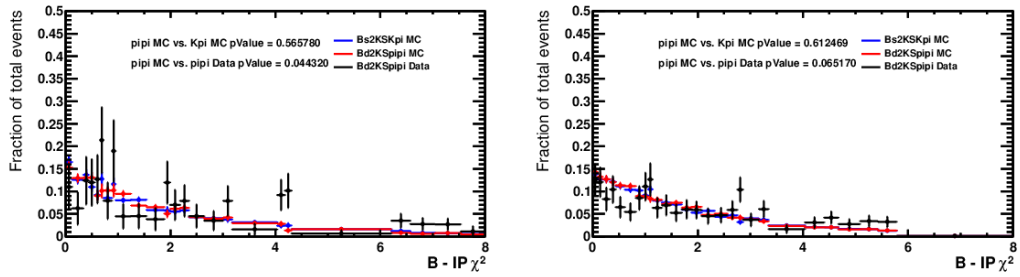

Figure 3.2: ${}_s\mathcal{P}lots$ of B impact parameter χ^2 distributions showing the $\mathrm{B}^0_s \to \mathrm{K}^0_\mathrm{S}\mathrm{K}^\pm\pi^\mp$ and $\mathrm{B}^0 \to \mathrm{K}^0_\mathrm{S}\pi^+\pi^-$ simulated events and $\mathrm{B}^0 \to \mathrm{K}^0_\mathrm{S}\pi^+\pi^-$ data events. Left: Long-Long K^0_S. Right: Down-Down K^0_S.

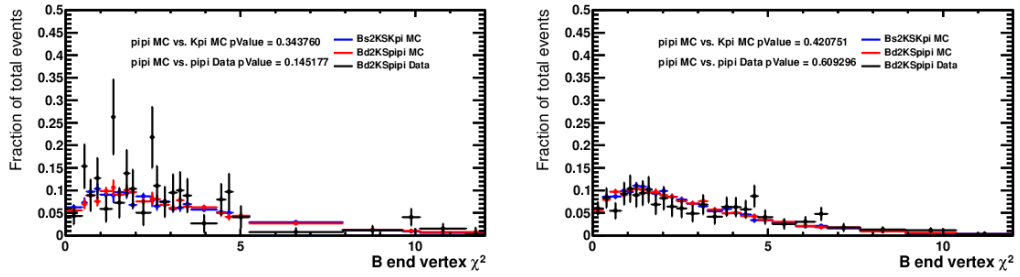

Figure 3.3: ${}_s\mathcal{P}lots$ of B end vertex χ^2 distributions showing the $\mathrm{B}^0_s \to \mathrm{K}^0_\mathrm{S}\mathrm{K}^\pm\pi^\mp$ and $\mathrm{B}^0 \to \mathrm{K}^0_\mathrm{S}\pi^+\pi^-$ simulated events and $\mathrm{B}^0 \to \mathrm{K}^0_\mathrm{S}\pi^+\pi^-$ data events. Left: Long-Long K^0_S. Right: Down-Down K^0_S.

The p-values shown suggest that there is negligible difference between the distributions of these variables between data and simulation and between the $\mathrm{K}^0_\mathrm{S}\pi^+\pi^-$ and $\mathrm{K}^0_\mathrm{S}\mathrm{K}^\pm\pi^\mp$ final states, except for the one value in Table 3.10. This p-value was recalculated with different histogram binning schemes and it was found to be an anomalous result, due to the particular binning scheme used. Therefore the decision to use the $\mathrm{B}^0 \to \mathrm{K}^0_\mathrm{S}\pi^+\pi^-$ simulation as the signal for training the neural net can be justified.

The sample of combinatorial background to train against was taken from the upper side-bands of the $\mathrm{K}^0_\mathrm{S}\pi^+\pi^-$ collision data. This is the same data as in Fig. 3.1 but with the requirement

$$5319.58 < m_B < 5429.58\,\mathrm{MeV}/c^2 \tag{3.1}$$

also applied to exclude the B^0 signal region and the partially reconstructed events from other background sources in the lower side-band. After training, the neural net

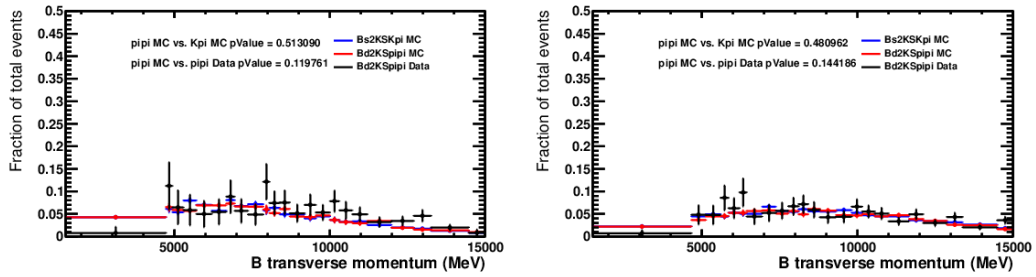

Figure 3.4: ${}_s\mathcal{P}lots$ of B transverse momentum distributions showing the $B^0_s \to K^0_S K^{\pm}\pi^{\mp}$ and $B^0 \to K^0_S\pi^+\pi^-$ simulated events and $B^0 \to K^0_S\pi^+\pi^-$ data events. Left: Long-Long K^0_S. Right: Down-Down K^0_S.

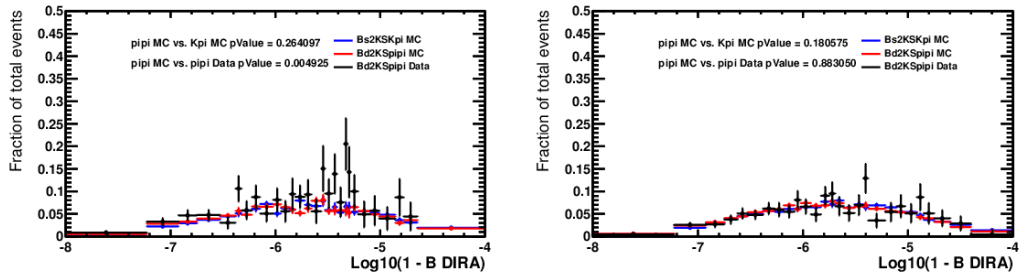

Figure 3.5: ${}_s\mathcal{P}lots$ of $\log_{10}$(1 - cosine of B pointing angle) distributions showing the $B^0_s \to K^0_S K^{\pm}\pi^{\mp}$ and $B^0 \to K^0_S\pi^+\pi^-$ simulated events and $B^0 \to K^0_S\pi^+\pi^-$ data events. Left: Long-Long K^0_S. Right: Down-Down K^0_S.

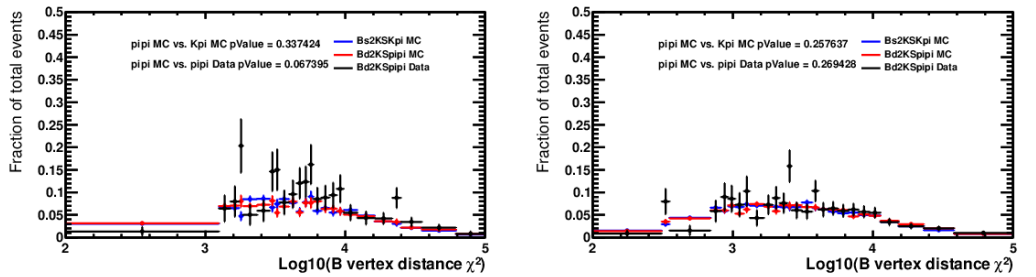

Figure 3.6: ${}_s\mathcal{P}lots$ of $\log_{10}$(B vertex distance χ^2) distributions showing the $B^0_s \to K^0_S K^{\pm}\pi^{\mp}$ and $B^0 \to K^0_S\pi^+\pi^-$ simulated events and $B^0 \to K^0_S\pi^+\pi^-$ data events. Left: Long-Long K^0_S. Right: Down-Down K^0_S.

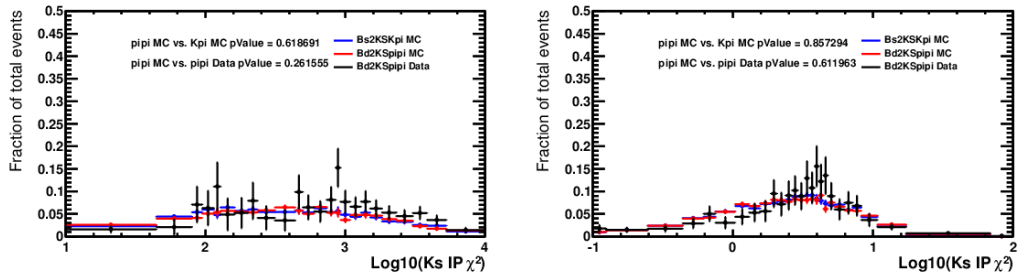

Figure 3.7: ${}_s\mathcal{P}lots$ of $\log_{10}(\mathrm{K^0_S}$ IP $\chi^2)$ distributions showing the $\mathrm{B^0_s} \to \mathrm{K^0_S K^\pm \pi^\mp}$ and $\mathrm{B^0} \to \mathrm{K^0_S \pi^+ \pi^-}$ simulated events and $\mathrm{B^0} \to \mathrm{K^0_S \pi^+ \pi^-}$ data events. Left: Long-Long $\mathrm{K^0_S}$. Right: Down-Down $\mathrm{K^0_S}$.

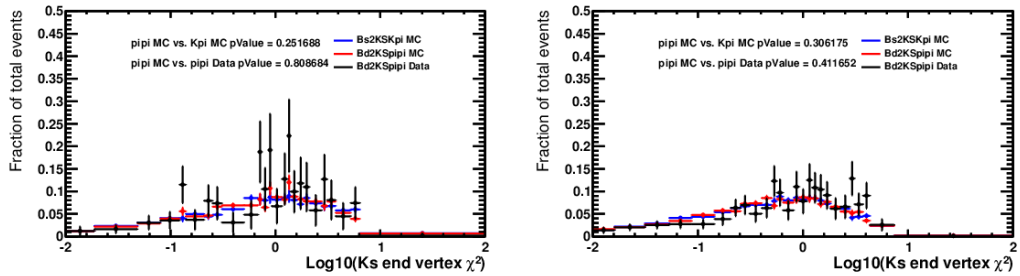

Figure 3.8: ${}_s\mathcal{P}lots$ of $\log_{10}(\mathrm{K^0_S}$ end vertex $\chi^2)$ distributions showing the $\mathrm{B^0_s} \to \mathrm{K^0_S K^\pm \pi^\mp}$ and $\mathrm{B^0} \to \mathrm{K^0_S \pi^+ \pi^-}$ simulated events and $\mathrm{B^0} \to \mathrm{K^0_S \pi^+ \pi^-}$ data events. Left: Long-Long $\mathrm{K^0_S}$. Right: Down-Down $\mathrm{K^0_S}$.

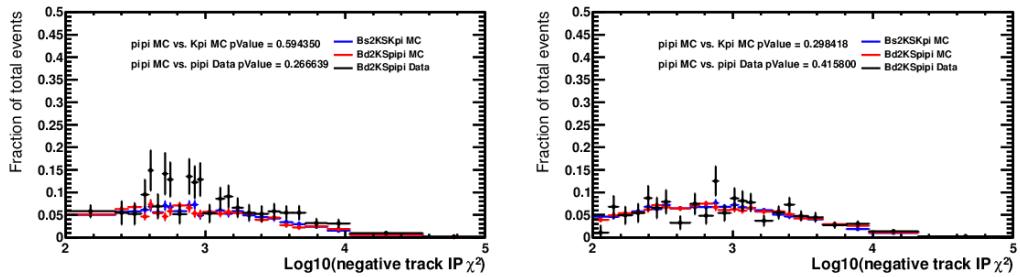

Figure 3.9: ${}_s\mathcal{P}lots$ of $\log_{10}(\text{negative track IP } \chi^2)$ distributions showing the $\mathrm{B^0_s} \to \mathrm{K^0_S K^\pm \pi^\mp}$ and $\mathrm{B^0} \to \mathrm{K^0_S \pi^+ \pi^-}$ simulated events and $\mathrm{B^0} \to \mathrm{K^0_S \pi^+ \pi^-}$ data events. Left: Long-Long $\mathrm{K^0_S}$. Right: Down-Down $\mathrm{K^0_S}$.

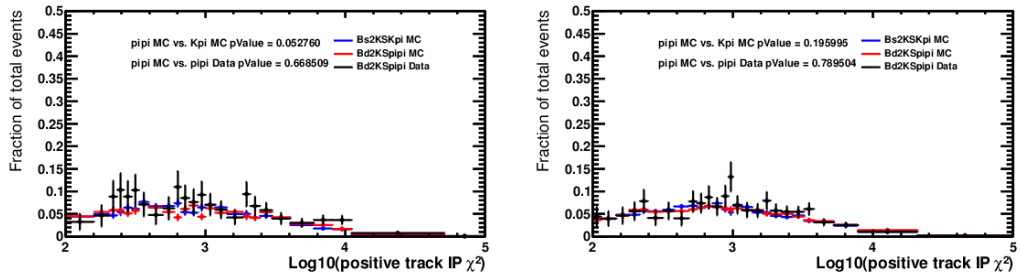

Figure 3.10: ${}_s\mathcal{P}lots$ of $\log_{10}$(positive track IP χ^2) distributions showing the $B^0_s \to K^0_S K^{\pm}\pi^{\mp}$ and $B^0 \to K^0_S \pi^+\pi^-$ simulated events and $B^0 \to K^0_S \pi^+\pi^-$ data events. Left: Long-Long K^0_S. Right: Down-Down K^0_S.

output values on the data samples are shown in Fig. 3.11. The neural net output distributions show that after training there is good separation between the signal events and the combinatorial data sample.

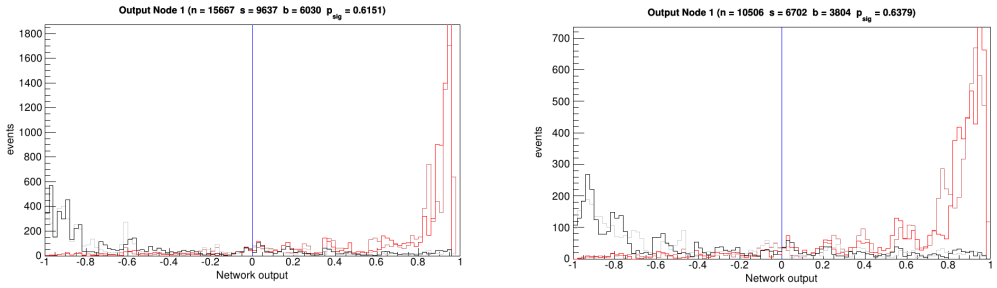

Figure 3.11: Neural net output on training data. Signal in red (dark red = boosted), combinatorial in black (grey = boosted). Left: Long-Long K^0_S sample. Right: Down-Down K^0_S sample.

Optimisation of neural net requirements

In order to choose the optimal cut value on the two neural nets the decision was made to maximize the figure of merit put forward by G. Punzi [62]. The figure of merit value, F_{Punzi}, is given by the equation

$$F_{\text{Punzi}} = \frac{\epsilon_{\text{Sig}}}{\frac{N_\sigma}{2} + \sqrt{N_{\text{Bkg}}}}. \tag{3.2}$$

Where N_σ is the required signal significance after the cut, N_{Bkg} is the remaining number of background events and ϵ_{Sig} is the efficiency of the cut on the signal events.

To evaluate the value of F_{Punzi} for various neural net cut values, data samples for the estimation of ϵ_{Sig} and N_{Bkg} are required. The signal data sample was taken from the $\mathrm{B}_s^0 \to \mathrm{K}_{\mathrm{S}}^0\mathrm{K}^{\pm}\pi^{\mp}$ simulation, with the same pre-selection requirements as during the training of the neural net. An additional requirement on the mass of the four-momentum combination of the $\mathrm{K}_{\mathrm{S}}^0$ and π tracks was made prior to the neural net requirement. This combination was mass constrained by the decay tree fitter so that the particle candidate masses are assumed to be the nominal values. This requirement was made in order to limit the signal events to only those events near to the $\mathrm{K}^{*\pm}$ mass region and was given by

$$791.66 < m_{\mathrm{K}_{\mathrm{S}}^0\pi} < 991.66\,\mathrm{MeV}/c^2. \tag{3.3}$$

The correlation of the B_s^0 mass with the neural net output value was also checked for these events. This is a useful check for bias in events which pass the final neural net cuts, as a significant correlation could indicate over training of the neural net. The 2D histograms of neural net value against the measured invariant B_s^0 mass combination and the correlation factors are given in Fig. 3.12. They show that the correlation factors are small and unlikely to indicate a significant bias in the neural net outs.

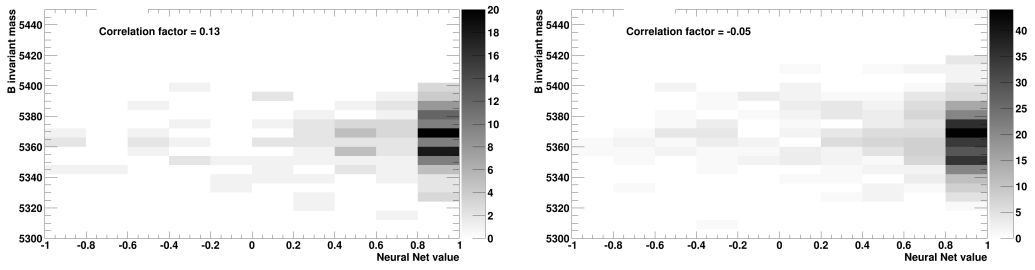

Figure 3.12: Correlation of neural net value against B_s^0 mass for $\mathrm{B}_s^0 \to \mathrm{K}_{\mathrm{S}}^0\mathrm{K}^{\pm}\pi^{\mp}$ simulated events. Left: Long-Long $\mathrm{K}_{\mathrm{S}}^0$ sample. Right: Down-Down $\mathrm{K}_{\mathrm{S}}^0$ sample.

For N_{Bkg}, the amount of combinatorial background under the B_s^0 mass peak and within the $\mathrm{K}^{*\pm}$ mass region in the $\mathrm{K}_{\mathrm{S}}^0\mathrm{K}^{\pm}\pi^{\mp}$ final state had to be estimated. In order to keep the signal region blind, the upper side-band of the $\mathrm{B}_s^0 \to \mathrm{K}_{\mathrm{S}}^0\mathrm{K}^{\pm}\pi^{\mp}$ region was used to estimate N_{Bkg} in the B_s^0 mass region. The upper side-band region was taken to be

$$5416.77 < m_B < 5491.77\,\mathrm{MeV}/c^2, \tag{3.4}$$

while the signal region was

$$5341.77 < m_B < 5391.77\,\mathrm{MeV}/c^2. \tag{3.5}$$

The upper side-band region was then fitted with a linear PDF, allowing the number of signal region events, prior to any neural net or $\mathrm{K}^{*\pm}$ mass requirements, to be extrapolated as shown in Fig. 3.13.

However, the number of combinatorial events after the requirement in Eq. 3.3 is relatively small, leading to an uncertain fit and estimate of N_{Bkg}. The ratio of the number of events in the upper side-band before and after the $\mathrm{K}^{*\pm}$ mass requirement and before any neural net requirements was found. The data sample prior to the $\mathrm{K}^{*\pm}$ mass requirement was used to fit the combinatorial background for a range of neural net cut values and give estimates of the combinatorial yield in the signal region. Each yield was then scaled down, using the previously measured ratio, to give an estimate of N_{Bkg} after the $\mathrm{K}^{*\pm}$ mass cut was applied.

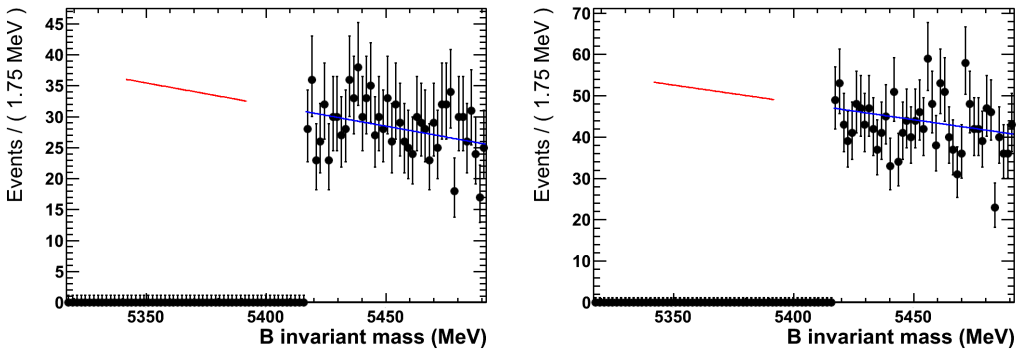

Figure 3.13: B invariant mass of $\mathrm{K}^0_{\mathrm{S}}\mathrm{K}^{\pm}\pi^{\mp}$ collision data. No neural net or $m_{\mathrm{K}^{*\pm}}$ requirements. Upper sideband modelled with a linear PDF (blue) and extrapolated under the B^0_s mass region (dashed red). Left: Long-Long $\mathrm{K}^0_{\mathrm{S}}$ sample. Right: Down-Down $\mathrm{K}^0_{\mathrm{S}}$ sample.

With $N_\sigma = 5$ and the values of ϵ_{Sig} and N_{Bkg} found by the methods outlined above, the values of F_{Punzi} were calculated for the range of neural net requirements. The values of F_{Punzi} and ϵ_{Sig} are plotted in Fig. 3.14. The optimal value of the neural net requirement was taken to be that which gives the maximum value of F_{Punzi}. The requirements chosen are listed in Table 3.11, which give signal efficiencies on $\mathrm{B}^0_s \to \mathrm{K}^0_{\mathrm{S}}\mathrm{K}^{\pm}\pi^{\mp}$ events of $\gtrsim 90\%$ and $\gtrsim 80\%$ for the Long-Long $\mathrm{K}^0_{\mathrm{S}}$ and Down-Down $\mathrm{K}^0_{\mathrm{S}}$ events respectively.

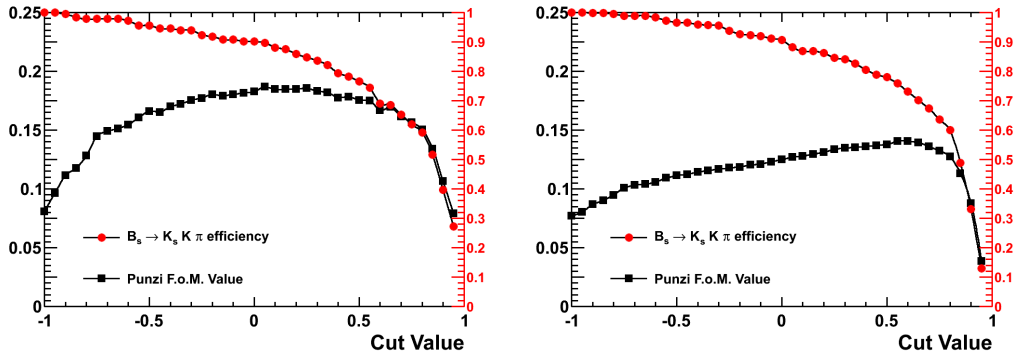

Figure 3.14: Punzi figure of merit value for neural network (black), using $N_\sigma = 5$. Also shown is the signal efficiency (red) at this cut. Left: Long-Long K^0_S sample. Right: Down-Down K^0_S sample.

Table 3.11: Neural net cuts applied to the different K^0_S types.

K^0_S type	Selection for neural net value
Long-Long K^0_S	$N > 0.05$
Down-Down K^0_S	$N > 0.55$

3.7 Particle identification

In this analysis the branching fractions were measured relative to the branching fraction of the $B^0 \to K^{*+}\pi^-$ decay, which has the final state $K^0_S\pi^+\pi^-$. The main signal decay being measured was $B^0_s \to K^{*\pm}K^{\mp}$, which has the final state $K^0_S K^{\pm}\pi^{\mp}$. To distinguish between these final states and those containing KK/pπ/*etc.*, PID cuts are applied to the charged tracks. In this analysis, the $DLL_{K\pi}$ and $DLL_{p\pi}$ delta log-likelihood variables produced by the RICH sub-detectors are used as the primary PID variables [63].

As in Section 3.6, the optimisation procedure followed the maximisation of a Punzi figure of merit. In order to find the value of ϵ_{Sig} for the evaluation of F_{Punzi}, from Eq. 3.2, a data sample of simulated $B^0_s \to K^{*\pm}K^{\mp}$ events was used. Due to differences between the simulation and data response of the RICH sub-detectors, the DLL variable distributions are not well modelled in the simulation. Therefore the well modelled p and p_T distributions were instead used in conjunction with the method of evaluating PID efficiencies described in Chapter 6, which utilizes calibration tracks of known PID from $D^{*+} \to D^0(K^-\pi^+)\pi^+$ decays. Prior to the efficiency

measurements, the simulated $B_s^0 \to K^{*\pm}K^{\mp}$ events had the full selection described in the previous sections applied, including the optimised neural net requirements and the $K^{*\pm}$ mass requirement from Eq. 3.3.

The most significant background that the $B_s^0 \to K^{*\pm}K^{\mp}$ PID requirements should reject is the $B^0 \to K^{*+}\pi^-$ decay mode, where the pion is mis-identified as a kaon. To find an estimate for $N_{\rm Bkg}$ for $B^0 \to K^{*+}\pi^-$ events in the $K_S^0K^{\pm}\pi^{\mp}$ final state, first a ratio of the measured values of the $B^0 \to K_S^0\pi^+\pi^-$ and $B^0 \to K^{*+}\pi^-$ branching fractions, given in Table 1.2, was made. This allows a simplistic estimation of $N_{\rm Bkg}$ for $B^0 \to K^{*+}\pi^-$ if the number of $B^0 \to K_S^0\pi^+\pi^-$ events reconstructed in the $K_S^0K^{\pm}\pi^{\mp}$ final state is known.

Since the $B_{d,s}^0 \to K_S^0h^{\pm}h'^{\mp}$ analysis [1] used the same data sample as this one as well as very similar selection requirements, the yield of $B^0 \to K_S^0\pi^+\pi^-$ events in the $K_S^0\pi^+\pi^-$ final state, prior to the PID requirements, should be approximately the same. This number of events was calculated by dividing the final yield of $B^0 \to K_S^0\pi^+\pi^-$ events in the previous analysis by the measured PID efficiency.

The p and p_T distributions for the simulated $B_s^0 \to K^{*\pm}K^{\mp}$ events should be similar to the distributions for the $B^0 \to K^{*+}\pi^-$ events. Therefore the $B_s^0 \to K^{*\pm}K^{\mp}$ simulated events were used with the procedure from Chapter 6 to calculate the $K^*\pi \to K^*K$ mis-ID efficiency. A range of mis-ID efficiencies were calculated for different $\mathrm{DLL}_{K\pi}$ requirements. Multiplying the mis-ID efficiency, ratio of branching fractions and yield of $B^0 \to K_S^0\pi^+\pi^-$ events calculated gave estimates for the remaining yield of $B^0 \to K^{*+}\pi^-$, $N_{\rm Bkg}$, after the different PID requirements are applied.

The PID requirements to reject proton tracks were taken from the $B_{d,s}^0 \to K_S^0h^{\pm}h'^{\mp}$ analysis and are shown in Table 3.12. As the background being rejected depends on separation of kaon and pion tracks, these proton PID requirements were held constant for the optimisation. The $\mathrm{DLL}_{K\pi}$ requirements were varied over the range $\mathrm{DLL}_{K\pi} < -5 \to 5$ for the pion tracks, and $\mathrm{DLL}_{K\pi} > 0 \to 10$ for the kaon tracks. As for the neural net optimisation $N_\sigma = 5$ for the evaluation of $F_{\rm Punzi}$. The values of the figure of merit are shown in Fig. 3.15.

In both the Long-Long K_S^0 and Down-Down K_S^0 cases, the value of $F_{\rm Punzi}$ is maximal when the pion $\mathrm{DLL}_{K\pi}$ track requirement is five. By fixing $\mathrm{DLL}_{K\pi} < 5$ for the pion tracks, the $F_{\rm Punzi}$ value and signal PID efficiency were plotted when only varying the kaon track requirement, see Fig. 3.16. As the signal efficiency drops with increasing values of $\mathrm{DLL}_{K\pi}$, the decision was made to take the kaon track requirement to be $\mathrm{DLL}_{K\pi} > 9$ for both the Long-Long K_S^0 and Down-Down K_S^0 events. These are not the optimal cuts, based on the potentially increasing value of

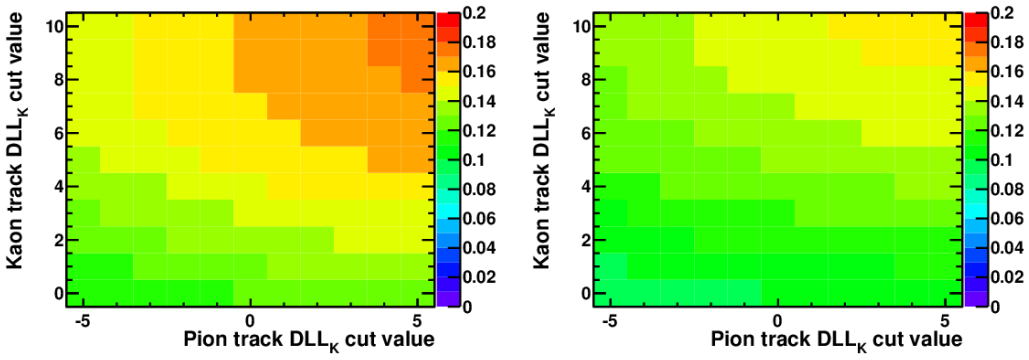

Figure 3.15: Punzi figure of merit value for various $\mathrm{DLL}_{\mathrm{K}\pi}$ cuts on the pion and kaon tracks. Efficiency from the $\mathrm{B}^0_s \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$ simulation, background estimate is of $\mathrm{B}^0 \to \mathrm{K}^{*+}\pi^-$ events. Left: Long-Long $\mathrm{K}^0_{\mathrm{S}}$ sample. Right: Long-Long $\mathrm{K}^0_{\mathrm{S}}$ sample

F_{Punzi} beyond the range of Fig. 3.16. However when performing the evaluation of the PID efficiency and the related systematics, see Section 6.4 and Section 7.3, it is necessary for the efficiency histograms to have a significant number of tracks in each bin after the relevant PID cuts are applied. Otherwise, the systematic uncertainty increases and the evaluation of that systematic is more difficult. Therefore the PID cuts were taken to be looser than the figure of merit curve alone would suggest is optimal, as making the cuts tighter than this began to adversely affect the efficiency histograms.

The data sample containing the $\mathrm{K}^0_{\mathrm{S}}\pi^+\pi^-$ final state was used to measure the yield of the normalisation mode $\mathrm{B}^0 \to \mathrm{K}^{*+}\pi^-$. To remove the cross-feed from $\mathrm{B}^0_s \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$ events $\mathrm{DLL}_{\mathrm{K}\pi}$ requirements were made to keep pion tracks and reject kaons. As either of the final state pion tracks could have come from the K^*, both of the $\mathrm{DLL}_{\mathrm{K}\pi}$ requirements should select against kaons equally. For simplicity, the $\mathrm{DLL}_{\mathrm{K}\pi}$ requirements were chosen so that the final PID efficiency on the $\mathrm{B}^0 \to \mathrm{K}^{*+}\pi^-$ decay would be similar to the efficiency for the $\mathrm{B}^0_s \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$. The full set of PID selection requirements are listed in Table 3.12.

3.8 Fitting ranges

The $\mathrm{B}^0_{d,s} \to \mathrm{K}^{*\pm}\mathrm{h}^{\mp}$ decay modes will peak in the m_{B} invariant mass and the $m_{\mathrm{K}^0_{\mathrm{S}}\pi}$ combination. The fits to the data sample, described in Chapter 5, use both of these variables as discriminating parameters. Therefore it was necessary to constrain the ranges of these observables to useful ranges.

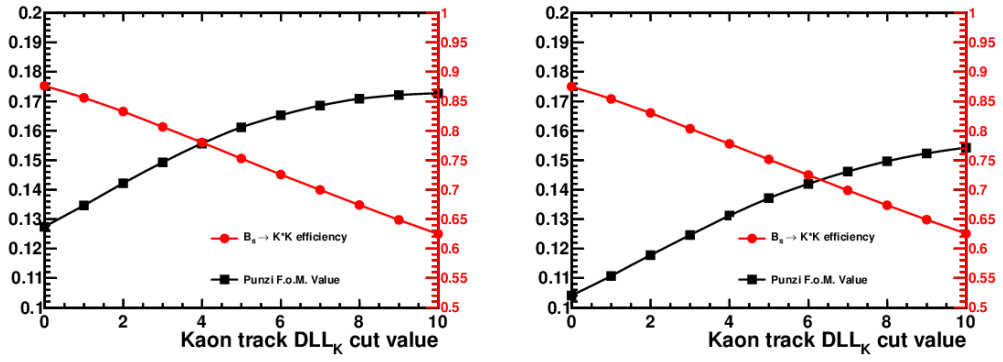

Figure 3.16: Punzi figure of merit value for various $\mathrm{DLL}_{\mathrm{K}\pi}$ cuts on the kaon tracks. Efficiency from the $\mathrm{B}^0_s \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$ simulation, background estimate is of $\mathrm{B}^0 \to \mathrm{K}^{*+}\pi^-$ events. Left: Long-Long $\mathrm{K}^0_{\mathrm{S}}$ sample. Right Down-Down $\mathrm{K}^0_{\mathrm{S}}$ sample.

Table 3.12: PID cuts applied to the particle hypothesis of a kaon or pion for the $\mathrm{K}^0_{\mathrm{S}}\pi^+\pi^-$ and $\mathrm{K}^0_{\mathrm{S}}\mathrm{K}^{\pm}\pi^{\mp}$ final states.

Track	π	K
$\mathrm{K}^0_{\mathrm{S}}\mathrm{K}^{\pm}\pi^{\mp}$ final state		
Kaon PID Requirement	$\mathrm{DLL}_{\mathrm{K}\pi} < 5$	$\mathrm{DLL}_{\mathrm{K}\pi} > 9$
Proton PID Requirement	$\mathrm{DLL}_{\mathrm{p}\pi} < 10$	$(\mathrm{DLL}_{\mathrm{p}\pi} - \mathrm{DLL}_{\mathrm{K}\pi}) < 10$
$\mathrm{K}^0_{\mathrm{S}}\pi^+\pi^-$ final state		
Kaon PID Requirement	$\mathrm{DLL}_{\mathrm{K}\pi} < -1$	-
Proton PID Requirement	$\mathrm{DLL}_{\mathrm{p}\pi} < 10$	-

The $m_{\rm B}$ range should fully contain both the $\rm B^0$ and $\rm B^0_s$ mass peaks. It should also be wide enough on the lower bound to include enough of the partially reconstructed backgrounds to fit. The upper bound should be high enough to include some combinatorial background to fit.

The $m_{\rm K^0_S\pi}$ observable range should be wide enough to contain the $\rm K^{*\pm}(892)$ resonance. However, the shape of the S-wave contribution may not be easily parameterised across a large range of $m_{\rm K^0_S\pi}$. Also the higher energy $\rm K^*$ resonances, specifically the $\rm K^*(1410)$, should be excluded. This implies that the range should be chosen carefully, to minimise the difficulties of constructing the fit model.

The final ranges chosen for the fit, which are applied as selection requirements, are shown in Table 3.13. In the case of the fit to the $\rm K^0_S K^\pm\pi^\mp$ final state sample, the $m_{\rm K^0_S\pi}$ variable was chosen to be the square root of the $\rm K^0_S\pi$ Dalitz plot variable, mass constrained using the decay tree fitter to the $\rm B^0_s \to K^0_S K^\pm\pi^\mp$ masses. The fit to the $\rm K^0_S\pi^+\pi^-$ final state sample uses the same $m_{\rm K^0_S\pi}$ Dalitz combination, but mass constrained to the $\rm B^0 \to K^0_S\pi^+\pi^-$ masses. The fitting variables and the distributions of the signal decays in them are discussed further in Chapter 5. Some general information on Dalitz variables is given in Appendix A.

Table 3.13: Fitting ranges of the $m_{\rm B}$ and $m_{\rm K^0_S\pi}$ variables.

Observable	Range (MeV)
$m_{\rm B}$	$5000 - 5500$
$m_{\rm K^0_S\pi}$	$650 - 1200$

4 Backgrounds

4.1 Introduction

There are a multitude of different potential backgrounds that could pass the current selection. It is therefore important to classify and quantify the dominant sources of background events, so that they may be accounted for in the fit model. This requires knowledge of the branching fractions of the background decays, efficiency for them to pass the selection requirements, as well as an understanding of the effects of mis-reconstruction. The background sources considered for the $K^0_S K^{\pm}\pi^{\mp}$ and $K^0_S \pi^+\pi^-$ final states are discussed in Sections 4.2-4.5.

This chapter does not contain any discussion of the fit model or the PDF shapes used to model the background shapes, these are described in Chapter 5. However, the shapes of the background PDFs used in the fit model have been plotted on the data in this chapter. This is to prevent the background shape plots being shown twice, once in this chapter without the PDFs and once in Chapter 5 with the PDF shapes. The fits of these PDFs to the data all had a fully converging fit and have a quoted estimated distance to the minimum log-likelihood, see Section 5.2. This distance to the minimum can be used as an estimate of the fit quality and comes from the fitting packages (MIGRAD and MINUIT). As a guide, lower suggest a better fit and values < 0.1 can be considered sufficiently well fitted. The plots

in this chapter are arbitrarily normalised, due to the re-weighting procedures, and should not be used as a definitive guide to the number of events in each data sample.

The estimated signal and background yields for the $K^0_S K^\pm \pi^\mp$ selection are described in Section 4.6. The decisions whether or not to include certain modes in the final fit model are discussed in Chapter 5.

4.2 Decays containing a real $K^{*\pm}$

Decays which contain real $K^{*\pm}$ particles will peak in the $m_{K^0_S\pi}$ spectrum with either the same, or a very similar shape as the signal decay modes. This means that they must either be reduced by the selection, or the fit model must use only the information from the B invariant mass spectrum to differentiate between the signal and these modes. Careful treatment of these modes is therefore essential and the modes considered are listed below.

The decay $B^0 \to K^{*+}\pi^-$

This is arguably the closest background mode to the $B^0_{d,s} \to K^{*\pm}K^\mp$ signal decays. It differs from the $B^0 \to K^{*\pm}K^\mp$ signal mode only by the misidentification of the charged pion track as a kaon. This causes the measured B invariant mass to be shifted upwards such that it lies under the $B^0_s \to K^{*\pm}K^\mp$ signal peak. However, this mode is greatly suppressed by the PID requirements described in Section 3.7 due to the optimisation strategy. Simulated $B^0 \to K^0_S\pi^+\pi^-$ events which were re-weighted by their Dalitz plot amplitudes to $B^0 \to K^{*+}\pi^-$ events are shown in Fig. 4.1. They are fitted with a Crystal Ball PDF [64] in the m_B spectrum and a spin-1 relativistic Breit-Wigner PDF in the $m_{K^0_S\pi}$ spectrum.

The decay $B^0_s \to K^{*\pm}K^\mp$

Although this mode is the main signal for this analysis in the $K^0_S K^\pm \pi^\mp$ final state, in the $K^0_S\pi^+\pi^-$ final state it enters as a potential background. The misidentification of the kaon track as a pion shifts the B invariant mass peak until it lies between the B^0 and B^0_s masses. As with the $B^0 \to K^{*+}\pi^-$ mode in the $K^0_S K^\pm \pi^\mp$ final state, this mode is suppressed by the PID requirements, with an expected misidentification rate of $\sim 2\%$. Simulated $B^0_s \to K^{*\pm}K^\mp$ events reconstructed in the $K^0_S\pi^+\pi^-$ final state, which were fitted with a Crystal-Ball PDF (m_B) and a spin-1 relativistic Breit-Wigner PDF ($m_{K^0_S\pi}$) are shown in Fig. 4.2.

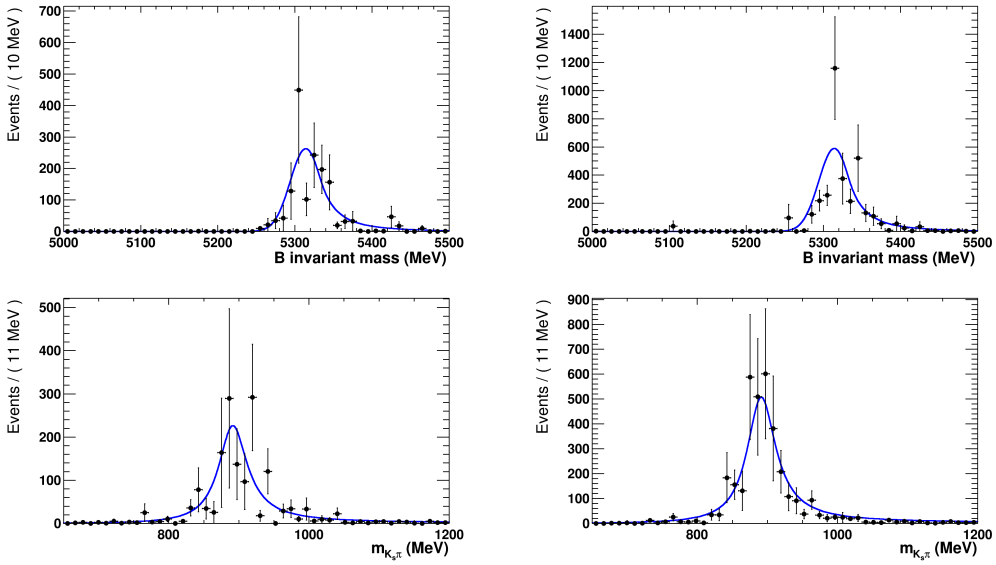

Figure 4.1: Background shapes from simulated events of $B^0 \to K^0_S\pi^+\pi^-$, reconstructed in the $K^0_S K^\pm\pi^\mp$ final state. Re-weighted as $B^0 \to K^{*+}\pi^-$ events using LAURA++. Estimated distance to minimum of fitted PDFs 1.9×10^{-6}. Top: m_B spectrum. Bottom: $m_{K^0_S\pi}$ spectrum. Left: Long-Long K^0_S sample. Right: Down-Down K^0_S sample.

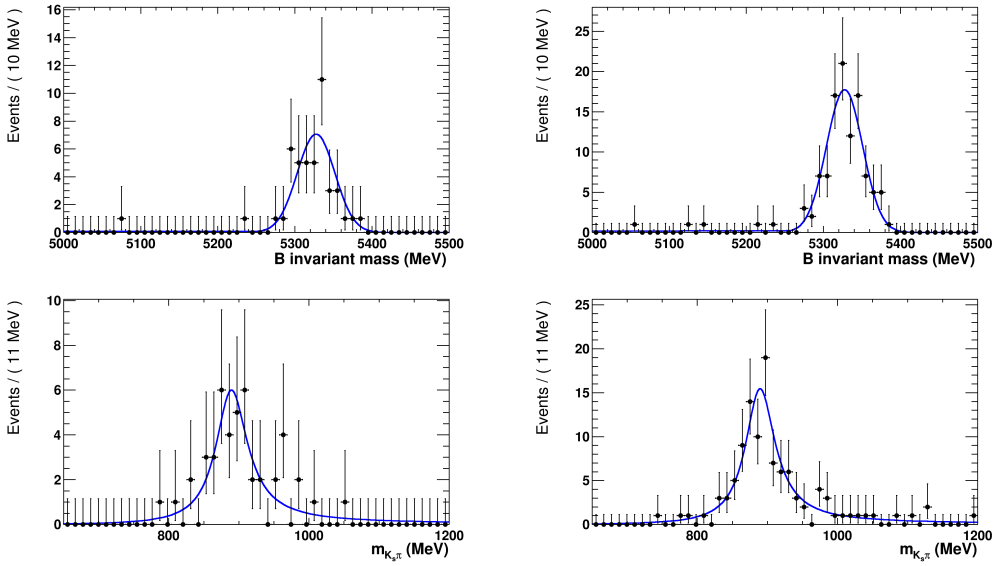

Figure 4.2: Background shapes from simulated events of $B_s^0 \to K^{*\pm}K^{\mp}$, reconstructed in the $K_S^0\pi^+\pi^-$ final state. Fitted with a Crystal Ball PDF (m_B) and a spin-1 relativistic Breit-Wigner PDF ($m_{K_S^0\pi}$). Estimated distance to minimum of fitted PDFs 1.6×10^{-6}. Top: m_B spectrum. Bottom: $m_{K_S^0\pi}$ spectrum. Left: Long-Long K_S^0 sample. Right: Down-Down K_S^0 sample.

The decay $\Lambda_b^0 \to \mathrm{K}^{*-}\mathrm{p}$

This mode is currently unmeasured, as is the non-resonant decay to the same final state. This makes the expected yield essentially unknowable and is estimated using assumptions about the similarity to the B meson decays which are similar at the quark level. However, the larger Λ_b mass is shifted such that it lies under the B_s^0 invariant mass region by the misidentification of the (anti-)proton as a kaon, allowing for events to potentially lie under the signal peak. As with the $\mathrm{B}^0 \to \mathrm{K}^{*+}\pi^-$ mode, the proton PID requirements should suppress this background. Simulated $\Lambda_b^0 \to \mathrm{K}_\mathrm{S}^0\pi^-\mathrm{p}$ events reconstructed in the $\mathrm{K}_\mathrm{S}^0\mathrm{K}^\pm\pi^\mp$ final state that were re-weighted by their Dalitz plot amplitudes to $\Lambda_b^0 \to \mathrm{K}^{*-}\mathrm{p}$ events are shown in Fig. 4.3. They are fitted with a RooKeysPdf in the m_B spectrum and a spin-1 relativistic Breit-Wigner PDF in the $m_{\mathrm{K}_\mathrm{S}^0\pi}$ spectrum. For a description of the RooKeysPdf, see Section 5.4.

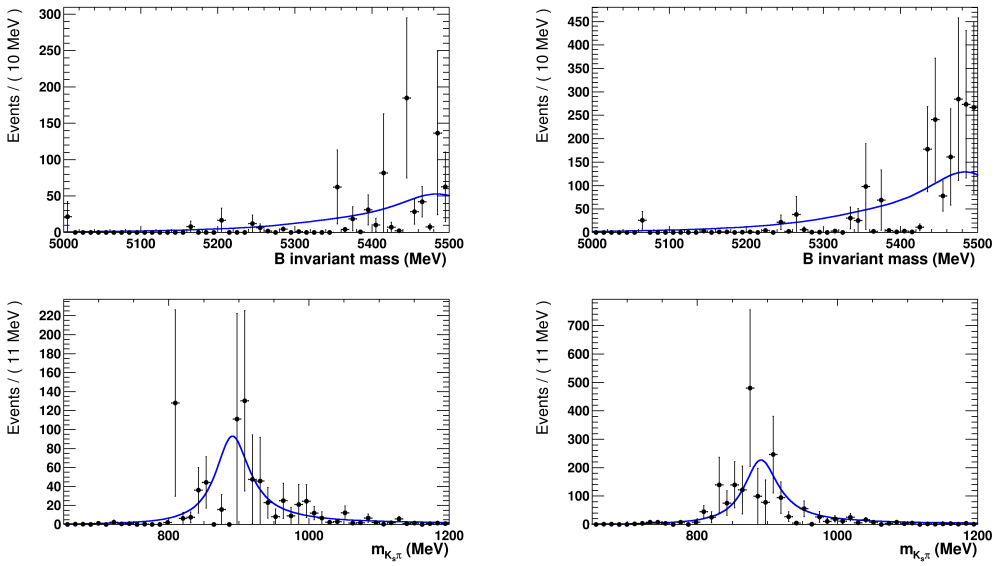

Figure 4.3: Background shapes from simulated events of $\Lambda_b^0 \to \mathrm{K}_\mathrm{S}^0\pi^-\mathrm{p}$, reconstructed in the $\mathrm{K}_\mathrm{S}^0\mathrm{K}^\pm\pi^\mp$ final state. Re-weighted as $\Lambda_b^0 \to \mathrm{K}^{*-}\mathrm{p}$ events using LAURA++. Fitted with a RooKeysPdf (m_B) and a spin-1 relativistic Breit-Wigner PDF ($m_{\mathrm{K}_\mathrm{S}^0\pi}$). Estimated distance to minimum of fitted PDFs 4.6×10^{-6}. Top: m_B spectrum. Bottom: $m_{\mathrm{K}_\mathrm{S}^0\pi}$ spectrum. Left: Long-Long K_S^0 sample. Right: Down-Down K_S^0 sample.

Partially reconstructed decays

Each of these modes contains an extra pion track which is not used to reconstruct the B vertex. For the $K^0_S K^{\pm}\pi^{\mp}$ final state one of the pions is additionally misidentified as a kaon, causing further suppression. The missing track causes the measured B invariant mass to be shifted to lower values, while the misidentification of a pion as a kaon has the opposite effect. This potentially allows B^0, B^+ and B^0_s background modes to stay in the signal region despite having a missing track.

In this analysis the decay modes $B^0 \to K^*(K^0_S\pi^+)\rho(\pi^-\pi^0)$, $B^+ \to K^*(K^0_S\pi^+)\pi^+\pi^-$ and $B^0_s \to K^*(K\pi^0)K^*(K^0_S\pi)$ are considered as potential partially reconstructed backgrounds for the $K^0_S K^{\pm}\pi^{\mp}$ final state. The shapes of the $B^0 \to K^*(K^0_S\pi^+)\rho(\pi^-\pi^0)$ and $B^0_s \to K^*(K\pi^0)K^*(K^0_S\pi)$ are shown in Figs. 4.4-4.5. The $B^+ \to K^*(K^0_S\pi^+)\pi^+\pi^-$ has a very similar shape to the $B^0 \to K^*(K^0_S\pi^+)\rho(\pi^-\pi^0)$ and the two modes are considered together for this analysis.

The $B^0 \to K^*(K^0_S\pi^+)\rho(\pi^-\pi^0)$ shape, when reconstructed in the $K^0_S\pi^+\pi^-$ final state, is shown in Fig. 4.6.

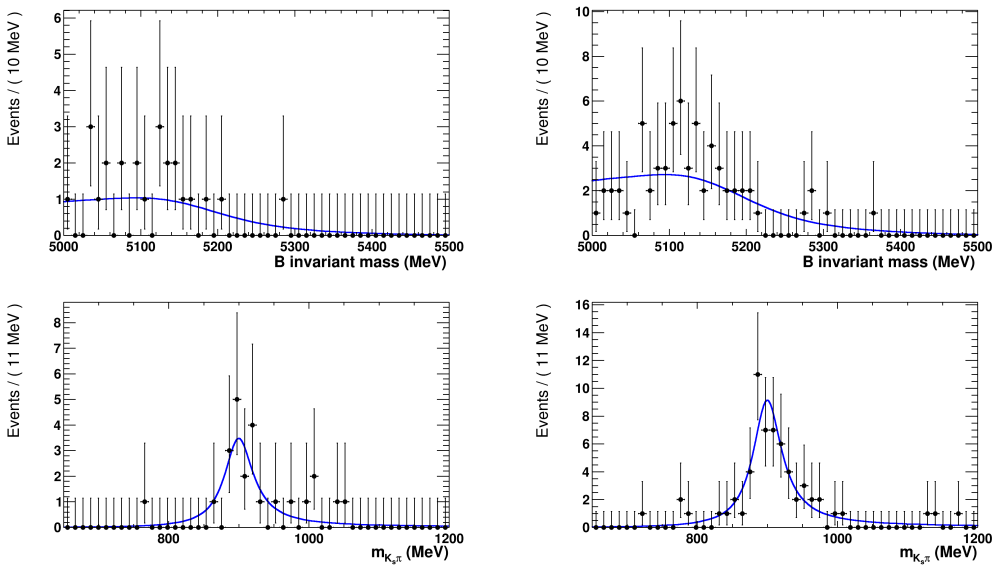

Figure 4.4: Background shapes from simulated events of $B^0 \to K^*(K^0_S\pi^+)\rho(\pi^-\pi^0)$, reconstructed in the $K^0_S K^{\pm}\pi^{\mp}$ final state. Fitted with a RooKeysPdf (m_B) and a spin-1 relativistic Breit-Wigner PDF ($m_{K^0_S\pi}$). Estimated distance to minimum of fitted PDFs 5.7×10^{-6}. Top: m_B spectrum. Bottom: $m_{K^0_S\pi}$ spectrum. Left: Long-Long K^0_S sample. Right: Down-Down K^0_S sample.

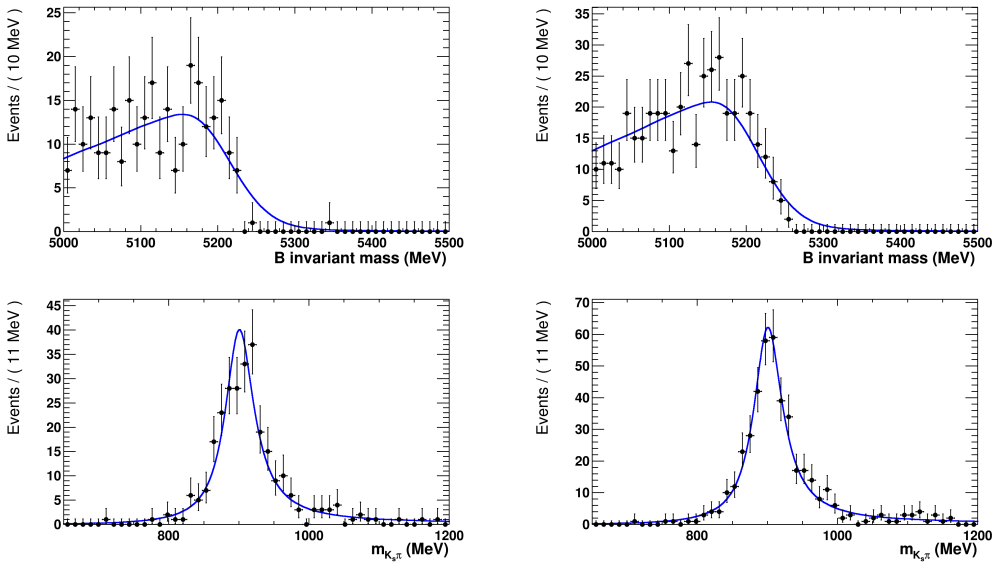

Figure 4.5: Background shapes from simulated events of $\mathrm{B}_s^0 \to \mathrm{K}^*(\mathrm{K}\pi^0)\mathrm{K}^*(\mathrm{K}_\mathrm{S}^0\pi)$. Fitted with a RooKeysPdf (m_B) and a spin-1 relativistic Breit-Wigner PDF ($m_{\mathrm{K}_\mathrm{S}^0\pi}$). Estimated distance to minimum of fitted PDFs 2.6×10^{-6}. Top: m_B spectrum. Bottom: $m_{\mathrm{K}_\mathrm{S}^0\pi}$ spectrum. Left: Long-Long K_S^0 sample. Right: Down-Down K_S^0 sample.

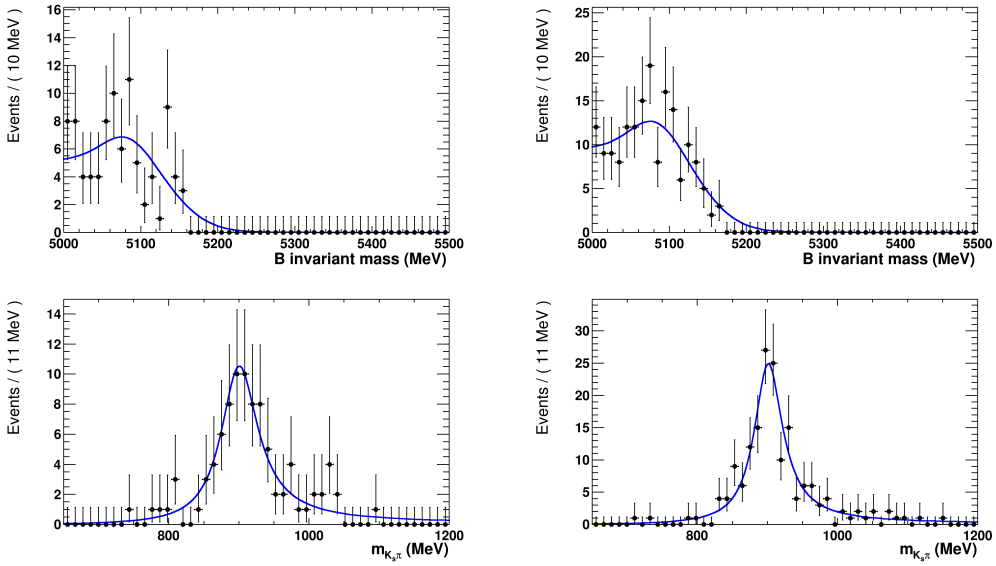

Figure 4.6: Background shapes from simulated events of $B^0 \to K^*(K^0_S\pi^+)\rho(\pi^-\pi^0)$, reconstructed in the $K^0_S\pi^+\pi^-$ final state. Fitted with a RooKeysPdf (m_B) and a spin-1 relativistic Breit-Wigner PDF ($m_{K^0_S\pi}$). Estimated distance to minimum of fitted PDFs 9.5×10^{-5}. Top: m_B spectrum. Bottom: $m_{K^0_S\pi}$ spectrum. Left: Long-Long K^0_S sample. Right: Down-Down K^0_S sample.

4.3 Non-resonant $K^0_S K^\pm \pi^\mp$ decays

When making the $B^0_{d,s} \to K^{*\pm} K^\mp$ selection, the final state $K^0_S K^\pm \pi^\mp$ track variables are used to select the signal decays. However, there is essentially no difference between the final state track variables of a $B^0_s \to K^{*\pm} K^\mp$ decay and any other $B^0_s \to K^0_S K^\pm \pi^\mp$ decay which proceeds via some other intermediate state. The same is true for the other $B^0_{d,s} \to K^{*\pm} \pi^\mp$ modes compared with the $B^0_{d,s} \to K^0_S \pi^+ \pi^-$ modes, in the $K^0_S \pi^+ \pi^-$ final state. Here we use non-resonant to refer to the multitude of intermediate states which decay into the correct final state, but do not contain a real $K^*(892)^\pm$. For example, the B^0_s could decay into $\overline{K}^*(1430)^0 K^0_S$ or $a_0(980)^+ \pi^-$ and the subsequent decays might give the correct final state.

These types of decay do not generally have a peaking resonant structure in the $K^{*\pm}$ mass region, however they will be indistinguishable from the signal decay contribution in the B invariant mass. There will certainly be a contribution to the final events from these non-resonant components and they must be separated by the fit model. The shapes of the non-resonant contributions in the $K^0_S K^\pm \pi^\mp$ final state were taken from the simulated events for the $B^0 \to K^0_S K^\pm \pi^\mp$ and $B^0_s \to K^0_S K^\pm \pi^\mp$, which were generated flat in the phase space of the decay. These shapes are shown in Figs. 4.7-4.8 and, as expected, the $m_{K^0_S \pi}$ spectra of the non-resonant modes do not contain any peaking components. Similarly, the $B^0 \to K^0_S \pi^+ \pi^-$ and $B^0_s \to K^0_S \pi^+ \pi^-$ phase space simulated events are shown in Figs. 4.9-4.10.

However, interference from these contributions with the $K^{*\pm}$ signal modifies both the signal and background PDF shapes, potentially with a peaking effect in the $m_{K^0_S \pi}$ spectrum. The systematic uncertainty that arises from neglecting this interference contribution is evaluated in Chapter 7.

4.4 Charmed backgrounds

Charmonium intermediate states

The decay $B^0 \to J/\psi\,(\mu^+\mu^-) K^0_S$ and those proceeding via higher mass charmonium states, such as $B^0 \to \psi(2S) K^0_S$, have topologically similar decays to the $K^0_S h^+ h^-$ decay modes. Due to the similar μ and π masses, the reconstruction of these decays does not shift the B invariant mass peak significantly. This allows these backgrounds to appear in the final selection. Fig. 4.11 shows the J/ψ and $\psi(2S)$ appearing in the $\pi^+\pi^-$ invariant mass for the selection used in the $B^0_{d,s} \to K^0_S h^\pm h'^\mp$ analysis. This background is removed via the muon ID cut described in Section 3.5.

There are also possible contributions from the hadronic decays of these char-

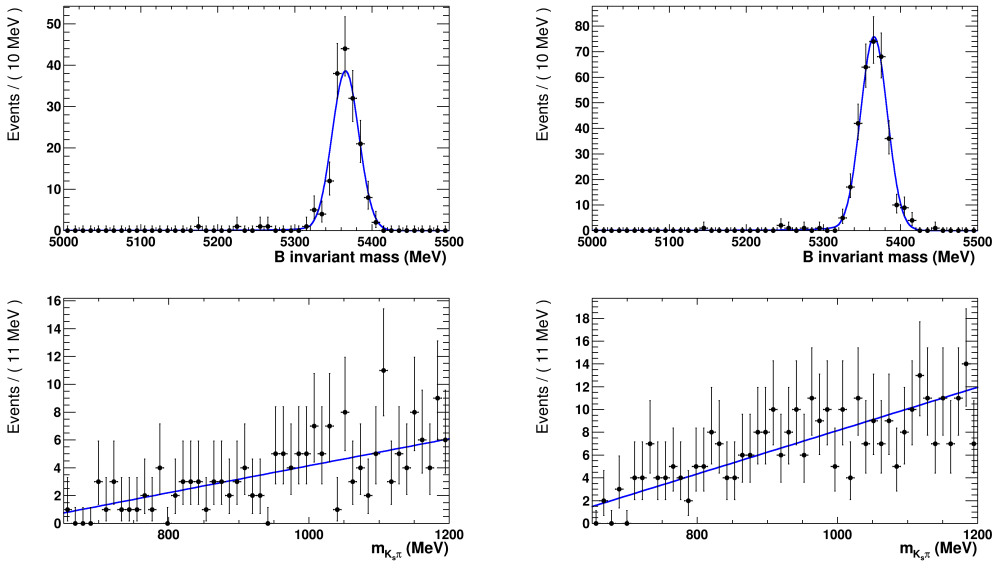

Figure 4.7: Background shapes from simulated events of $B_s^0 \to K_S^0 K^\pm \pi^\mp$. Fitted with a Crystal Ball PDF (m_B) and a linear PDF ($m_{K_S^0\pi}$). Estimated distance to minimum of fitted PDFs 7.3×10^{-5}. Top: m_B spectrum. Bottom: $m_{K_S^0\pi}$ spectrum. Left: Long-Long K_S^0 sample. Right: Down-Down K_S^0 sample.

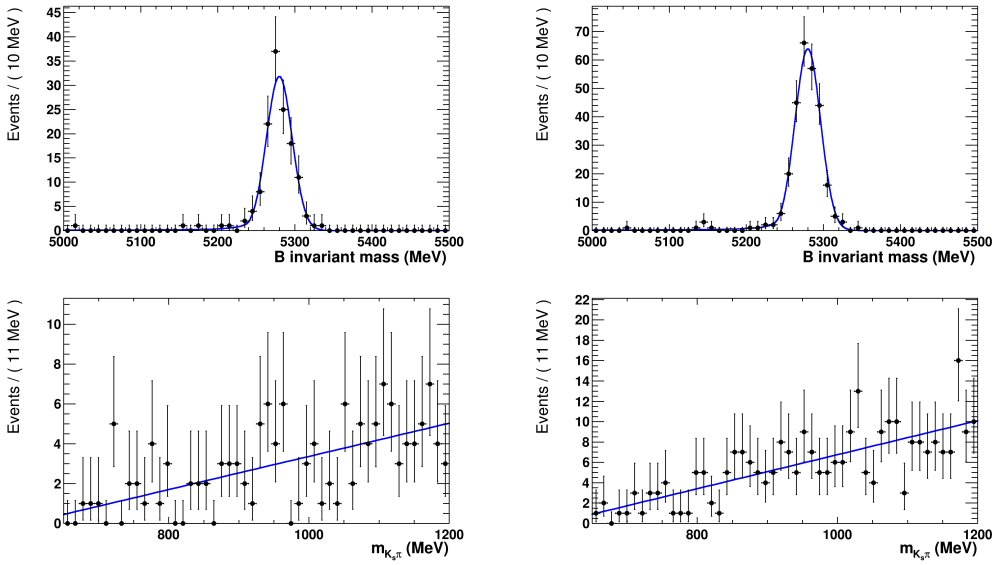

Figure 4.8: Background shapes from simulated events of $B^0 \to K_S^0 K^\pm \pi^\mp$. Fitted with a Crystal Ball PDF (m_B) and a linear PDF ($m_{K_S^0\pi}$). Estimated distance to minimum of fitted PDFs 4.5×10^{-6}. Top: m_B spectrum. Bottom: $m_{K_S^0\pi}$ spectrum. Left: Long-Long K_S^0 sample. Right: Down-Down K_S^0 sample.

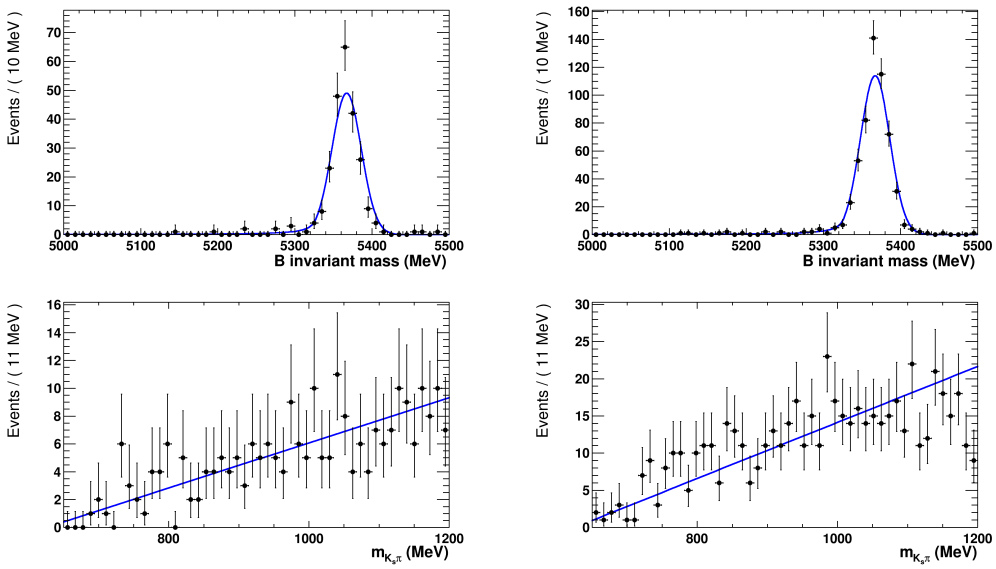

Figure 4.9: Background shapes from simulated events of $B_s^0 \to K_S^0\pi^+\pi^-$. Fitted with a Crystal Ball PDF (m_B) and a linear PDF ($m_{K_S^0\pi}$). Estimated distance to minimum of fitted PDFs 1.3×10^{-5}. Top: m_B spectrum. Bottom: $m_{K_S^0\pi}$ spectrum. Left: Long-Long K_S^0 sample. Right: Down-Down K_S^0 sample.

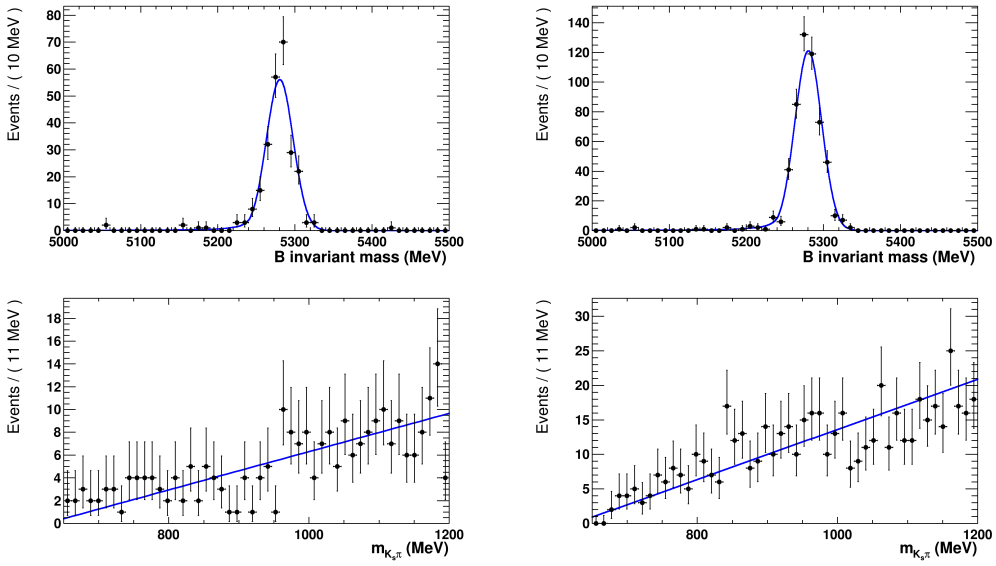

Figure 4.10: Background shapes from simulated events of $B^0 \to K_S^0\pi^+\pi^-$. Fitted with a Crystal Ball PDF (m_B) and a linear PDF ($m_{K_S^0\pi}$). Estimated distance to minimum of fitted PDFs 1.0×10^{-5}. Top: m_B spectrum. Bottom: $m_{K_S^0\pi}$ spectrum. Left: Long-Long K_S^0 sample. Right: Down-Down K_S^0 sample.

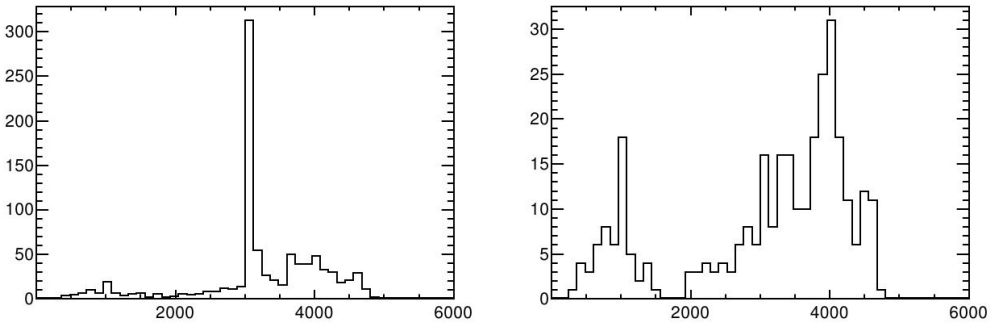

Figure 4.11: Mass combination of the $\pi^+\pi^-$ for $K^0_S\pi^+\pi^-$ Down-Down K^0_S combinations consistent with the B^0 hypothesis using 2011 data. Left: Before `isMuon` cut. Right: After `isMuon` cut. [1]

monium states. As described in Section 3.5 these events are removed by vetoing the relevant $\pi^+\pi^-$ and K^+K^- mass combinations.

Baryonic charmed decays

If a proton from a Λ_b decay is misidentified as a π or K it will shift the reconstructed mass into the B^0_s mass region. When directly selecting for proton tracks, Fig. 4.12 shows a clear Λ_c resonance in the $K^0_S\overline{p}$ mass combination in a subset of the 2011 data sample.

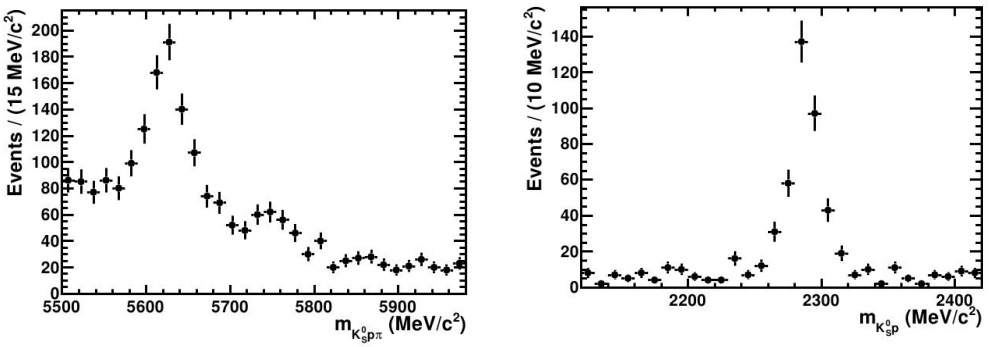

Figure 4.12: Mass combinations of $K^0_S\pi^+\overline{p}$ and $K^0_S\overline{p}$ from 2011 Down-Down K^0_S data showing baryonic peaks. Left: Mass combination of the $K^0_S\pi^+\overline{p}$, showing a Λ_b peak. Right: Mass combinations of the $K^0_S\overline{p}$ for $K^0_S\pi^+\overline{p}$ combinations consistent with Λ_b hypothesis, showing a Λ_c peak. [1]

The mass combinations $K^0_S\pi$ and $K^0_S K$ were used to search for final state $D^-_s p$ decays. Both the Cabibbo favoured $D^-_s \to K^0_S K^-$ and Cabibbo suppressed $D^-_s \to K^0_S \pi^-$ decays were considered.

After the applied proton PID requirements and charmed mass vetoes were applied, see Section 3.5, a negligible number of events for these baryonic modes remained.

Partially reconstructed decays

Charmed modes exist where the D meson decays to three tracks, one of which is not reconstructed. Due to the missing track, these cannot be efficiently vetoed and an additional misidentification of a pion track as a kaon could lead to events from these modes being in the signal region.

The two modes considered for the $K^0_S K^\pm \pi^\mp$ final state are the $B^+ \to D^0(K^0_S\pi^+\pi^-)K^+$ and $B^+ \to D^0(K^0_S\pi^+\pi^-)\pi^+$ decay modes. They have similar PDF shapes and together they are abbreviated as the $B^+ \to D^0 h$ modes. Although the $B^+ \to D^0(K^0_S\pi^+\pi^-)\pi^+$ mode has a much larger branching fraction than $B^+ \to D^0(K^0_S\pi^+\pi^-)K^+$, the additional suppression by the kaon PID requirement makes the expected yields comparable, see Tables 4.5-4.6. The $B^+ \to D^0(K^0_S\pi^+\pi^-)K^+$ shapes from simulated events, reconstructed in the $K^0_S K^\pm \pi^\mp$ final state, are shown in Fig. 4.13.

For the $K^0_S\pi^+\pi^-$ final state, the $B^+ \to D^0(K^0_S\pi^+\pi^-)\pi^+$ is dominant due to the suppression of $B^+ \to D^0(K^0_S\pi^+\pi^-)K^+$ by both the PID requirements and the branching fraction. The shapes of the $B^+ \to D^0(K^0_S\pi^+\pi^-)\pi^+$ from simulated events, reconstructed in the $K^0_S\pi^+\pi^-$ final state, are shown in Fig. 4.14.

In order to more easily see the shapes and positions of the various signal and background decay contributions. Fig. 4.15 shows arbitrarily normalised PDFs in the m_B spectrum for the Down-Down K^0_S $K^0_S K^\pm \pi^\mp$ final state. Here it is easy to see that the cross feed from the $B^0 \to K^{*+}\pi^-$ decays will lie under the signal peaks. Additionally, if there are large contributions from the $\Lambda^0_b \to K^{*-}p$ or partially reconstructed decays the tails of these distributions may lie under the signal region. Not shown are the $K^0_S\pi^+\pi^-$ and Long-Long K^0_S PDFs as they have very similar distributions. However, in addition to these decays a major contribution to the background events in the signal region is likely to be the combinatorial background.

4.5 Combinatorial background

Combinatorial background arises due to the reconstruction of a vertex from unrelated tracks. Both the stripping selection and multivariate selection are designed

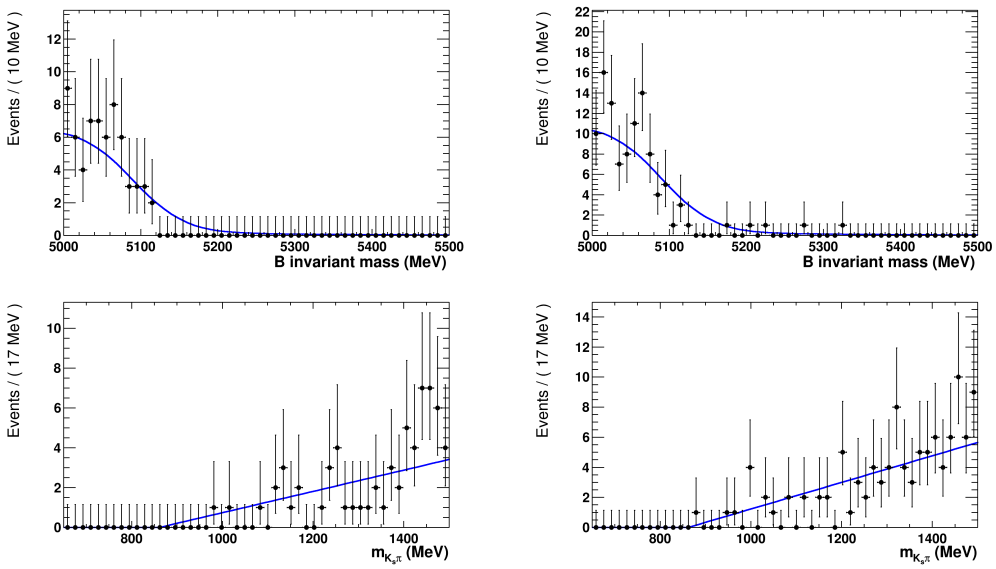

Figure 4.13: Background shapes from simulated events of $B^+ \to D^0(K^0_S\pi^+\pi^-)K^+$, reconstructed in the $K^0_S K^\pm\pi^\mp$ final state. Fitted with a RooKeysPdf (m_B) and a linear PDF ($m_{K^0_S\pi}$). Estimated distance to minimum of fitted PDFs 3.5×10^{-4}. Top: m_B spectrum. Bottom: $m_{K^0_S\pi}$ spectrum. Left: Long-Long K^0_S sample. Right: Down-Down K^0_S sample.

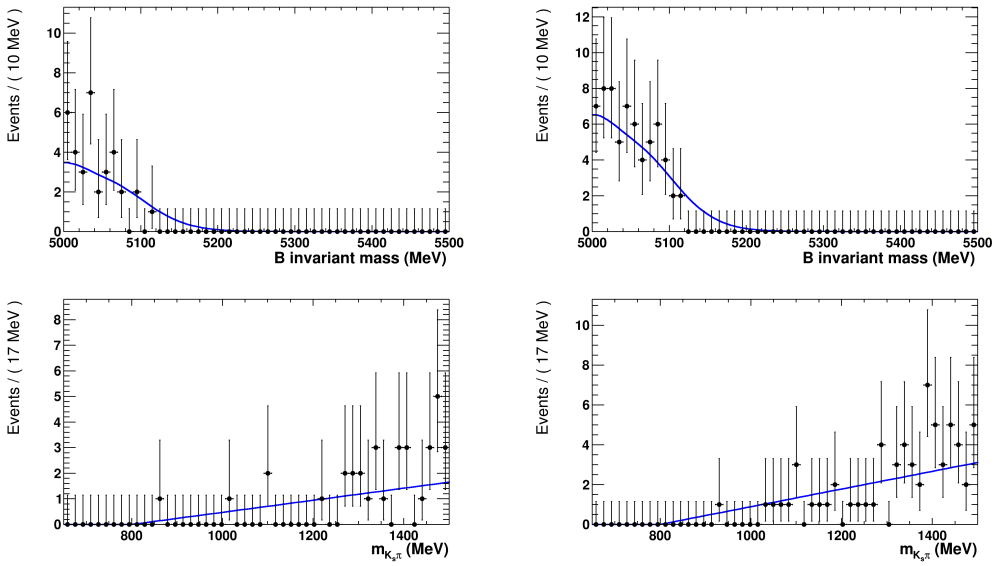

Figure 4.14: Background shapes from simulated events of $B^+ \to D^0(K^0_S\pi^+\pi^-)\pi^+$, reconstructed in the $K^0_S\pi^+\pi^-$ final state. Fitted with a RooKeysPdf (m_B) and a linear PDF ($m_{K^0_S\pi}$). Estimated distance to minimum of fitted PDFs 5.1×10^{-8}. Top: m_B spectrum. Bottom: $m_{K^0_S\pi}$ spectrum. Left: Long-Long K^0_S sample. Right: Down-Down K^0_S sample.

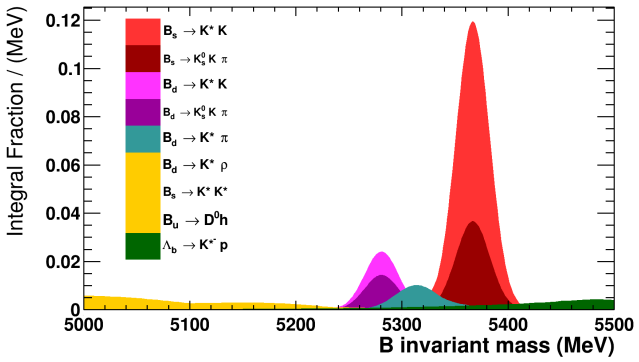

Figure 4.15: Background m_B shapes for multiple B and Λ_b signal and background decays from simulated events reconstructed in the Down-Down K^0_S $K^0_S K^\pm\pi^\mp$ final state. The $B^0_s \to K^{*\pm}K^\mp$ peak is normalised to unity. The rest of the decays are normalised to arbitrary values to allow easier visualisation of the peaks and their estimated contributions. The partially reconstructed decays are all given the same colour as they have similar distributions in the m_B spectrum.

to quickly and efficiently remove candidates of this type. However, some amount of combinatorial background will always pass these selections. Combinatorial background does not arise from a real decay, but is instead due to the random combination of tracks that do not necessarily come from the same particle decay. As such, it does not peak in either the m_{B} or $m_{\mathrm{K}^0_{\mathrm{S}}\pi}$ mass spectra. This makes it much easier to discriminate between this and signal, even with relatively large amounts of background events.

The upper m_{B} side-band events are used to obtain an estimate of the combinatorial shapes, as shown in Figs. 4.16-4.17. Here it can be seen that there is very little combinatorial background in the upper side-band of the Down-Down $\mathrm{K}^0_{\mathrm{S}}$ events, owing to the tighter neural net selection.

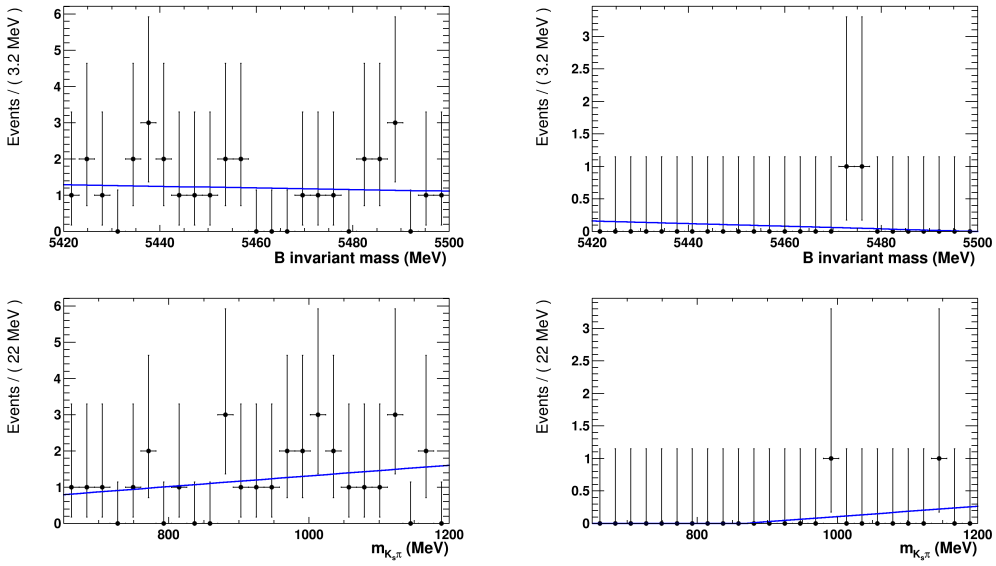

Figure 4.16: Upper m_{B} side-band events in the $\mathrm{K}^0_{\mathrm{S}}\mathrm{K}^{\pm}\pi^{\mp}$ final state. Used to estimate the combinatorial background shapes which are fitted with linear PDFs in both spectra. Estimated distance to minimum of fitted PDFs 5.3×10^{-6}. Top: m_{B} spectrum. Bottom: $m_{\mathrm{K}^0_{\mathrm{S}}\pi}$ spectrum. Left: Long-Long $\mathrm{K}^0_{\mathrm{S}}$ sample. Right: Down-Down $\mathrm{K}^0_{\mathrm{S}}$ sample.

4.6 Estimated yields

Signal yields

In order to estimate the number of signal events the equation

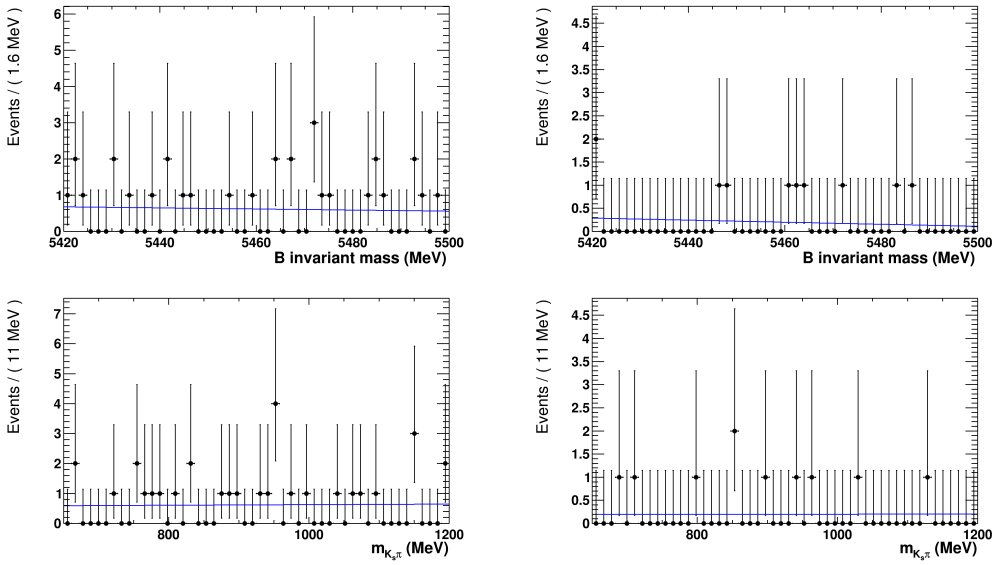

Figure 4.17: Upper m_{B} side-band events in the $\mathrm{K}^0_{\mathrm{S}}\pi^+\pi^-$ final state. Used to estimate the combinatorial background shapes which are fitted with linear PDFs in both spectra. Estimated distance to minimum of fitted PDFs 1.3×10^{-6}. Top: m_{B} spectrum. Bottom: $m_{\mathrm{K}^0_{\mathrm{S}}\pi}$ spectrum. Left: Long-Long $\mathrm{K}^0_{\mathrm{S}}$ sample. Right: Down-Down $\mathrm{K}^0_{\mathrm{S}}$ sample.

$$N_{\rm Sig} = 2\,\mathcal{L}\,\sigma_{\rm b\overline{b}}\,f_s\,\mathcal{B}\,({\rm Sig})\;\varepsilon^{\rm Acc}_{\rm Sig}\,\varepsilon^{\rm Sel}_{\rm Sig}, \tag{4.1}$$

was used. Here $\mathcal{L}$ is the integrated luminosity of the data sample used, which for this analysis was $\int \mathcal{L} = 1.01 \pm 0.03$ fb^{-1} [1]. $\sigma_{\rm b\overline{b}}$ is the b$\overline{\rm b}$ production cross-section within the LHCb acceptance, which has been measured to be $\sigma_{\rm b\overline{b}} = 75.3 \pm 5.4 \pm 13.0\,\mu$b [65], where the first error is statistical and the second is systematic. The product $2\,\mathcal{L}\,\sigma_{\rm b\overline{b}}$ then gives the number of b quarks produced, which will hadronise to form mesons and baryons. The hadronisation fractions of b quarks are listed in Table 4.1.

Table 4.1: Average b-hadron production fractions, averaged from LEP, Tevatron and LHCb data [34].

Quantity		Hadronisation fraction
$\rm B^0$ or $\rm B^+$	$f_{d,u}$	0.401 ± 0.007
$\rm B^0_s$	f_s	0.107 ± 0.005
$\rm B^0_s/B^0$ ratio	f_s/f_d	0.266 ± 0.015
Λ_b	f_{baryon}	0.091 ± 0.015

The detector acceptance efficiency for decays to $\rm K^0_S K^\pm \pi^\mp$ final states was measured previously to be $\varepsilon^{\rm Acc}_{\rm K^0_S K^\pm \pi^\mp} = 19.3 \pm 0.1\%$ [1]. This can be used as a good estimate of $\varepsilon^{\rm Acc}_{\rm Sig}$ as the final state particles are the same. The branching fractions $\mathcal{B}\,({\rm Sig})$ have not yet been measured. Theoretical predictions for the branching fractions are given in Table 4.2. The value, $\varepsilon^{\rm Sel}_{\rm Sig}$ is the selection efficiency for the signal events. This was measured by applying the selection to the simulated events and the results are given in Table 4.3. Using Eq. 4.1 the estimated signal yields were calculated and the results are given in Table 4.4.

Table 4.2: Theoretical predictions for *CP* averaged branching fractions of $\rm K^{*\pm}K^\mp$ signal decay modes. Calculated from Tables 1.3-1.4.

Quantity	B.F. ($\times 10^{-6}$)
$\rm B^0_s \to K^{*\pm}K^\mp$	14 ± 2
$\rm B^0 \to K^{*\pm}K^\mp$	0.15 ± 0.04

Background yields

The yields of the various background modes can also be estimated. The equation

Table 4.3: Naive total selection efficiency of signal modes from simulated events.

Decay mode	Long-Long K^0_S efficiency (%)	Down-Down K^0_S efficiency (%)
$B^0_s \to K^{*\pm}K^{\mp}$	0.117 ± 0.002	0.228 ± 0.003
$B^0 \to K^{*\pm}K^{\mp}$	0.111 ± 0.002	0.226 ± 0.003

Table 4.4: Estimated signal yields for $B^0_s \to K^{*\pm}K^{\mp}$ and $B^0 \to K^{*\pm}K^{\mp}$ modes.

Decay mode	Long-Long K^0_S yields	Down-Down K^0_S yields
$B^0_s \to K^{*\pm}K^{\mp}$	51 ± 9	99 ± 18
$B^0 \to K^{*\pm}K^{\mp}$	2.0 ± 0.5	4 ± 1

$$R_{\rm Bkg} = \frac{N_{\rm Bkg}}{N_{B^0_s \to K^{*\pm}K^{\mp}}} = \frac{f_\alpha}{f_s} \frac{\mathcal{B}\,({\rm Bkg})}{\mathcal{B}\,(B^0_s \to K^{*\pm}K^{\mp})} \frac{\varepsilon^{\rm Sel}_{\rm Bkg}}{\varepsilon^{\rm Sel}_{B^0_s \to K^{*\pm}K^{\mp}}} \frac{\varepsilon^{\rm Acc}_{\rm Bkg}}{\varepsilon^{\rm Acc}_{\rm Sig}}, \tag{4.2}$$

gives the formula for calculating the ratio of expected background events to the $B^0_s \to K^{*\pm}K^{\mp}$ events. In this equation, f_α is the hadronisation fraction for the relevant background mode. The acceptance efficiencies of the background modes are taken from the simulation. The branching fractions, selection efficiencies and relative yields of the various background modes are given in Tables 4.5-4.6. The relative combinatorial background yield can not estimated by this method, as it arises from random combinations of tracks.

Table 4.5: Branching fraction values used in the calculation of the background yield estimates. Unmeasured modes are estimated by comparison to similar measured modes.

Decay mode	B.F. ($\times 10^{-6}$)	Reference
$B^0 \to K^{*+}\pi^-$	8.5 ± 0.7	[39]
$\Lambda_b^0 \to K^{*-}p$	8.5 ± 1.0	-
$B^0 \to K^*(K_S^0\pi^+)\rho(\pi^-\pi^0)$	10.3 ± 2.6	[34]
$B^+ \to K^*(K_S^0\pi^+)\pi^+\pi^-$	75.3 ± 10.1	[34]
$B_s^0 \to K^*(K\pi^0)K^*(K_S^0\pi)$	5.1 ± 1.4	-
$B^0 \to K_S^0K^\pm\pi^\mp$ (NR)	1.9 ± 0.4	-
$B_s^0 \to K_S^0K^\pm\pi^\mp$ (NR)	28.7 ± 5.6	-
$B^+ \to D^0K^+$	383 ± 45	[34]
$B^+ \to D^0\pi^+$	4750 ± 190	[34]
$D^0 \to K_S^0\pi^+\pi^-$	28200 ± 1900	[4]
$B^+ \to D^0(K_S^0\pi^+\pi^-)K^+$	10.8 ± 1.5	-
$B^+ \to D^0(K_S^0\pi^+\pi^-)\pi^+$	134 ± 11	-

Table 4.6: Estimated background yields, relative to the $B_s^0 \to K^*K$ mode.

Decay mode	Long-Long K_S^0 relative yield	Down-Down K_S^0 relative yield
$B^0 \to K^{*+}\pi^-$	0.03 ± 0.01	0.02 ± 0.01
$\Lambda_b^0 \to K^{*-}p$	0.05 ± 0.01	0.01 ± 0.01
$B^0 \to K^*(K_S^0\pi^+)\rho(\pi^-\pi^0)$	0.004 ± 0.003	0.004 ± 0.002
$B^+ \to K^*(K_S^0\pi^+)\pi^+\pi^-$	0.03 ± 0.02	0.03 ± 0.01
$B_s^0 \to K^*(K\pi^0)K^*(K_S^0\pi)$	0.04 ± 0.01	0.03 ± 0.01
$B^0 \to K_S^0K^\pm\pi^\mp$ (NR)	0.05 ± 0.01	0.05 ± 0.01
$B_s^0 \to K_S^0K^\pm\pi^\mp$ (NR)	0.25 ± 0.06	0.25 ± 0.05
$B^+ \to D^0K^+$	0.01 ± 0.01	0.01 ± 0.01
$B^+ \to D^0\pi^+$	0.03 ± 0.02	0.01 ± 0.01

5 Fit model

5.1 Introduction

The selection requirements made are designed to minimise the number of background events and maximise the number of signal events. However the remaining events will certainly not be pure signal, indeed there may not be any signal events in the final data sample at all. Therefore it is necessary to estimate the number of signal and background events in the sample without bias, under various assumptions of the level of signal yield.

There are many standard fitting techniques and computer packages designed for this type of analysis. A brief overview of those used in this analysis is given in Section 5.2. The fit model used to describe the signal and background shapes is detailed in Sections 5.3-5.4. An important part of validating the fit model is the simulation studies, which can reveal stability issues and bias within the fit model. These studies are described in Section 5.5.

5.2 Fitting techniques

In particle physics we are often attempting to produce a model to represent the underlying physical processes being investigated. The parameters of the model can

either be set to the previously measured values, or estimated by fitting the model to the observed data. The standard way to fit to data is an *Extended Maximum Likelihood Fit.*

Maximum likelihood fits

The Likelihood ($\mathcal{L}$) is defined as

$$\mathcal{L}(\boldsymbol{\alpha}) = \prod_{i}^{N} [P(\boldsymbol{x}_i, \boldsymbol{\alpha})], \tag{5.1}$$

where N is the number of events and $P(\boldsymbol{x}_i, \boldsymbol{\alpha})$ is a Probability Density Function (PDF) [66]. The PDF depends on both a set of parameters, $\boldsymbol{\alpha}$, and a set of observables, $\boldsymbol{x}_i$, whose values may vary with each event, i. By varying the values of the parameters of the PDF the likelihood can be maximised, giving the best estimate for the true values of the parameters.

As the likelihood can be calculated for each event, there is no need to group the events into bins as in χ^2 fitting. Unbinned methods are favourable as they use the full information available from the events, whereas binning the data reduces the available information. The choice of binning by the user can also introduce additional bias in the fit.

Extended fits

It is, in general, the case that the ensemble of events consists of both signal and background contributions. Therefore it is useful to make the full PDF a sum of independent signal and background PDFs. For a single signal and background contribution, the normalised PDF can be written as

$$P(\boldsymbol{x}, \boldsymbol{\alpha}_{\text{sig}}, \boldsymbol{\alpha}_{\text{bkg}}) = f_{\text{sig}} P_{\text{sig}}(\boldsymbol{x}, \boldsymbol{\alpha}_{\text{sig}}) + (1 - f_{\text{sig}}) P_{\text{bkg}}(\boldsymbol{x}, \boldsymbol{\alpha}_{\text{bkg}}), \tag{5.2}$$

where f_{sig} is the fraction of signal events in the total. P_{sig} and P_{bkg} are the individually normalised signal and background PDFs respectively.

However if the absolute amounts of signal and background events are of interest, then the PDF can be *Extended*. This refers to the inclusion of a Poisson distribution term with a mean, ν, equal to the total of the contributing yields, $\nu = N_{\text{sig}} + N_{\text{bkg}}$. With the inclusion of this distribution, allowing separate yield parameters for each contribution, the extended likelihood equation becomes

$$\mathcal{L}(\nu, \boldsymbol{\alpha}) = \frac{e^{-\nu}}{N!} \prod_{i}^{N} [\nu P(\boldsymbol{x}_i, \boldsymbol{\alpha})]. \tag{5.3}$$

Pull distributions

As fit models and the parameter estimation algorithms become increasingly complicated it becomes necessary to use diagnostic tools to check for problems. A frequently used method is to perform an ensemble of fits to toy Monte Carlo events generated from the PDFs. Then the pull distributions of the fitted parameters are plotted [67]. The pull value for a parameter, α, for experiment, i, is given by

$$g_i = \frac{\alpha_i^{fit} - \alpha_i^{gen}}{\sigma_i^{\alpha_{fit}}}, \tag{5.4}$$

where α_i^{fit} is the fitted value of the parameter α, $\sigma_i^{\alpha_{fit}}$ is its corresponding uncertainty, and α_i^{gen} is the value used to generate the toy Monte Carlo sample. The value $\alpha_i^{fit} - \alpha_i^{gen}$ is known as the residual, r_i. As the ensemble size increases, the distribution of g_i should approach a normal distribution of mean zero and unit width if the underlying fit model is unbiased and stable.

5.3 $\mathrm{B}^0_{d,s} \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$ Signal model

As described in Chapter 4, both The $\mathrm{B}^0_{d,s} \to \mathrm{K}^0_{\mathrm{S}}\mathrm{K}^{\pm}\pi^{\mp}$ background modes and the signal $\mathrm{B}^0_{d,s} \to \mathrm{K}^*\mathrm{K}$ modes will peak in the B invariant mass variable, $m_{\mathrm{K}^0_{\mathrm{S}}\mathrm{K}\pi}$, with the same PDF shapes. It is therefore practically impossible to extract the number of $\mathrm{B}^0_{d,s} \to \mathrm{K}^*\mathrm{K}$ events when fitting only in $m_{\mathrm{K}^0_{\mathrm{S}}\mathrm{K}\pi}$. However, the $\mathrm{B}^0_{d,s} \to \mathrm{K}^*\mathrm{K}$ contains a $\mathrm{K}^{*\pm}$ resonance, which will peak in the $m_{\mathrm{K}^0_{\mathrm{S}}\pi}$ invariant mass while the $\mathrm{B}^0_{d,s} \to \mathrm{K}^0_{\mathrm{S}}\mathrm{K}^{\pm}\pi^{\mp}$ will not. In this analysis a two-dimensional (2D) fit is used in the $m_{\mathrm{K}^0_{\mathrm{S}}\mathrm{K}\pi}$ and $m_{\mathrm{K}^0_{\mathrm{S}}\pi}$ invariant mass variables. The ranges of these variables are given in Table 3.13. It should be noted that to improve the resolution of the $\mathrm{K}^{*\pm}$ resonance, the $m_{\mathrm{K}^0_{\mathrm{S}}\pi}$ mass variable is always calculated using a vertex fit where an additional constraint is applied that the invariant mass of the $\mathrm{K}^0_{\mathrm{S}}\mathrm{K}^{\pm}\pi^{\mp}$ combination must be that of the nominal B^0_s mass.

Additionally, the different $\mathrm{K}^0_{\mathrm{S}}$ samples gave two distinct data samples to fit. However, it would be better to combine the information from both data samples, in order to get a more precise fit. To achieve this, a simultaneous fit was performed to both data samples at once. The final fit model contains certain shared parameters between the separate $\mathrm{K}^0_{\mathrm{S}}$ models, whose values are measured by the fit to the $\mathrm{K}^0_{\mathrm{S}}$

samples simultaneously. The descriptions of the final fit model parameters are given in Table 5.3.

When creating a 2D fit model, there is the possibility for the observables to be correlated. If they are, the fit model should reflect this by including correlation terms in the PDF. To ascertain if there is correlation between the fit variables, they are plotted using simulated signal events, passing the full selection, in Fig. 5.1. The correlation coefficients for these plots are reproduced in Table 5.1 and show that there is no significant correlation. Correlation terms between the two observables of the fit are therefore not included in the fit model.

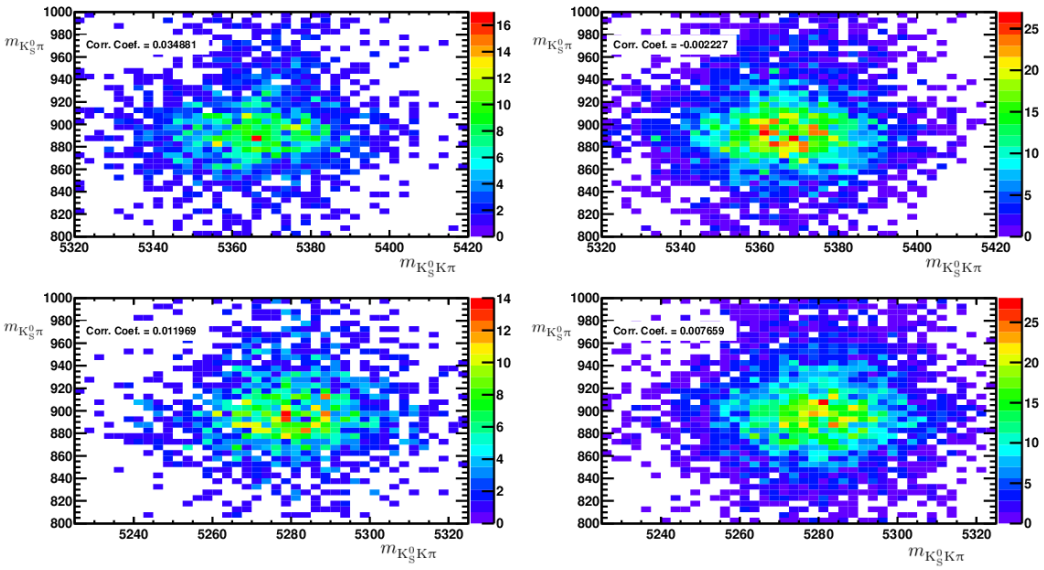

Figure 5.1: Correlation plots between the fit variables in simulated $B^0_{d,s} \to K^*K$ decays. Top: $B^0_s \to K^*K$. Bottom: $B^0 \to K^*K$. Left: Long-Long K^0_S sample. Right: Down-Down K^0_S sample.

Table 5.1: Correlation coefficients between the $m_{K^0_SK\pi}$ and $m_{K^0_S\pi}$ fit variables for simulated signal events.

Mode	Correlation coefficient
$B^0_s \to K^*K$ (Long-Long K^0_S)	0.035
$B^0_s \to K^*K$ (Down-Down K^0_S)	-0.002
$B^0 \to K^*K$ (Long-Long K^0_S)	0.012
$B^0 \to K^*K$ (Down-Down K^0_S)	0.008

The signal peak in the $m_{\mathrm{K^0_S K\pi}}$ variable is parameterised by a Crystal Ball PDF, a Gaussian peak with an exponential tail on one side [64]. To reduce the number of free parameters, the exponential tail parameters of the Crystal Ball functions are fixed from fits to the simulated events. The signal peak in the $m_{\mathrm{K^0_S \pi}}$ variable is modelled with a spin-1 Relativistic Breit-Wigner PDF.

However, the yield of $\mathrm{B^0 \to K^{*\pm} K^{\mp}}$ events is expected to be very small, see Table 4.4. After testing the fit to toy Monte Carlo, generated with the expected signal yields, it was shown that the $\mathrm{B^0 \to K^{*\pm} K^{\mp}}$ PDF was unstable. To correct this, the means and widths of the $\mathrm{B^0 \to K^{*\pm} K^{\mp}}$ peaks are parameterised in terms of the $\mathrm{B^0_s \to K^{*\pm} K^{\mp}}$ PDF values. The mean of the $\mathrm{B^0 \to K^{*\pm} K^{\mp}}$ Crystal Ball peak, $\bar{m}^{\mathrm{B^0}}_{m_{\mathrm{K^0_S K\pi}}}$, is then given by the relation

$$\bar{m}^{\mathrm{B^0}}_{m_{\mathrm{K^0_S K\pi}}} = \bar{m}^{\mathrm{B^0_s}}_{m_{\mathrm{K^0_S K\pi}}} - \Delta\bar{m}_{\mathrm{B}}, \tag{5.5}$$

where $\Delta\bar{m}_{\mathrm{B}}$ is the difference between the $\mathrm{B^0_s}$ and $\mathrm{B^0}$ masses.

A similar relation is also used for the mean parameter of the $\mathrm{B^0 \to K^{*\pm} K^{\mp}}$ Breit-Wigner PDF, $\bar{m}^{\mathrm{B^0}}_{m_{\mathrm{K^0_S \pi}}}$, using the difference $\Delta\bar{m}_{\mathrm{K^*}}$. This value may not be zero due to the use of the Decay Tree Fitter to constrain the $\mathrm{K^0_S K^{\pm} \pi^{\mp}}$ combination for this variable, to the nominal $\mathrm{B^0_s}$ mass. This can cause the $\mathrm{K^*}$ peak from the $\mathrm{B^0}$ signal decay to have a different mean or width to the $\mathrm{K^*}$ coming from a true $\mathrm{B^0_s}$ decay.

The widths of the $\mathrm{B^0 \to K^{*\pm} K^{\mp}}$ mass peaks in the Crystal Ball and Breit-Wigner PDFs are also parameterised by the complementary variables from the $\mathrm{B^0_s \to K^{*\pm} K^{\mp}}$ PDFs. Rather than the difference between the widths, the ratio of widths was used. So the $\mathrm{B^0 \to K^{*\pm} K^{\mp}}$ Crystal Ball width, $\sigma^{\mathrm{B^0}}_{m_{\mathrm{K^0_S K\pi}}}$, is tied to the width of the $\mathrm{B^0_s \to K^{*\pm} K^{\mp}}$ Crystal Ball peak by the relation

$$\sigma^{\mathrm{B^0}}_{m_{\mathrm{K^0_S K\pi}}} = \sigma^{\mathrm{B^0_s}}_{m_{\mathrm{K^0_S K\pi}}} . R_{\mathrm{B}} \tag{5.6}$$

where R_{B} is the ratio of the $\mathrm{B^0 \to K^{*\pm} K^{\mp}}$ Crystal Ball width, to that of the $\mathrm{B^0_s \to K^{*\pm} K^{\mp}}$. Similarly, the width of the $\mathrm{B^0 \to K^{*\pm} K^{\mp}}$ $\mathrm{K^*}$ peak is parameterised by the one in the $\mathrm{B^0_s \to K^{*\pm} K^{\mp}}$ PDF.

The values of the R and Δ parameters are fixed in the fit and the values taken from the fits to the simulated events in the relevant decay modes. While measurements exist for these parameters, there may be unexpected differences in the values due to reconstruction effects. For example, the $m_{\mathrm{K^0_S \pi}}$ variable comes from a vertex fit which assumes that the $\mathrm{K^0_S K^{\pm} \pi^{\mp}}$ combination forms the $\mathrm{B^0_s}$ mass.

While in the $B^0 \to K^{*\pm}K^{\mp}$ decays they instead form the B^0 mass, therefore the differences in the means and widths of the B^0 and B_s^0 peaks may not conform to the nominal values. Therefore the simulated events which have the same reconstruction requirements applied are used to fix these parameters.

For the final fit, each of the free mean and width shape parameters has a Gaussian constraint applied to them. The width of the constraint was taken from the uncertainty on the parameter during a fit to the simulated events. After fitting to $B^0_{d,s} \to K^*K$ simulated events, the projected fits are plotted in Figs. 5.2-5.3. These show that for both signal modes, these PDFs fit the simulated data shapes well.

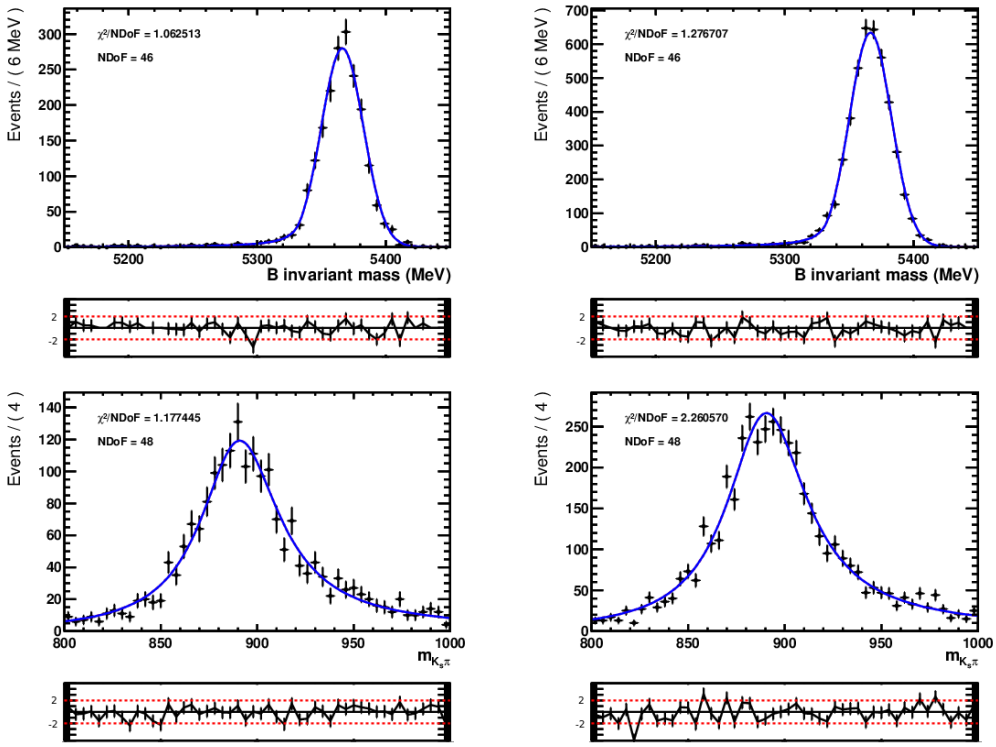

Figure 5.2: $B_s^0 \to K^*K$ signal fit model (blue) fitted to simulated events (black). The 2D fit model is projected into the individual $m_{K_S^0K\pi}$ and $m_{K_S^0\pi}$ observables with the pull distributions shown underneath. Top: $m_{K_S^0K\pi}$. Bottom: $m_{K_S^0\pi}$. Left: Long-Long K_S^0 sample. Right: Down-Down K_S^0 sample.

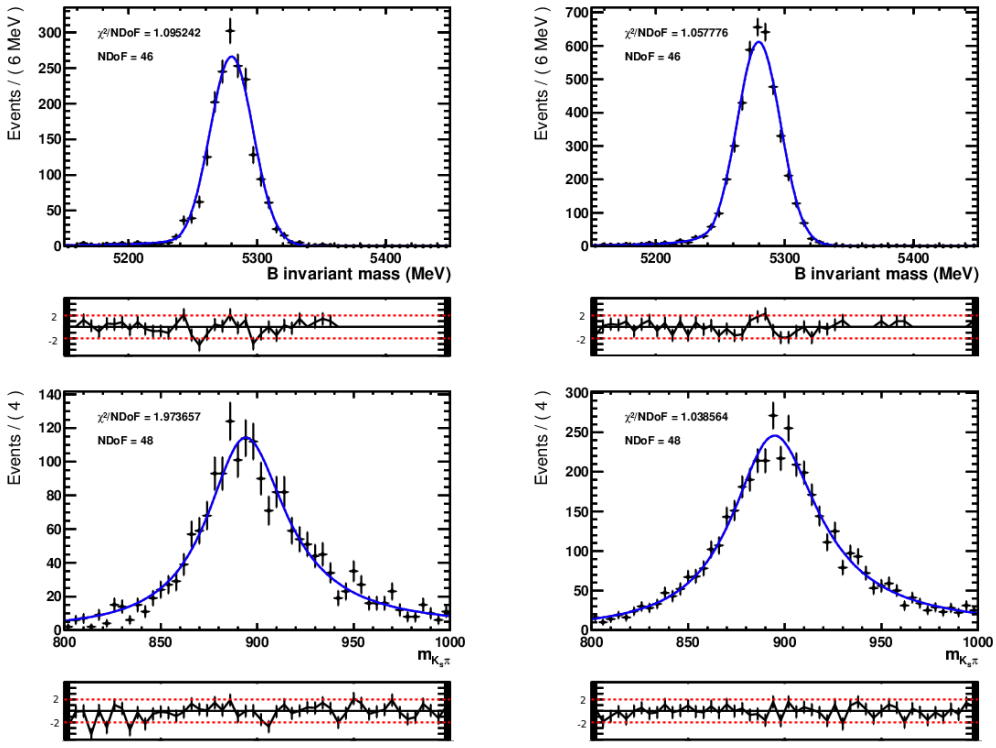

Figure 5.3: $B^0 \to K^*K$ signal fit model fitted to simulated events. The 2D fit model is projected into the individual $m_{K^0_S K\pi}$ and $m_{K^0_S \pi}$ observables with the pull distributions shown underneath. Top: $m_{K^0_S K\pi}$. Bottom: $m_{K^0_S \pi}$. Left: Long-Long K^0_S sample. Right: Down-Down K^0_S sample.

5.4 $B^0_{d,s} \to K^{*\pm}K^{\mp}$ Background models

The background fit model is the sum of various background components. The component PDFs used are:

- $B^0 \to K^{*+}\pi^-$.
- $\Lambda^0_b \to K^{*-}p$.
- $B^0 \to K^*(K^0_S\pi^+)\rho(\pi^-\pi^0)$ and $B^+ \to K^*(K^0_S\pi^+)\pi^+\pi^-$ combined into one PDF, abbreviated as $B \to K^*\pi\pi$
- $B^0_s \to K^*(K\pi^0)K^*(K^0_S\pi)$.
- $B^0_s \to K^0_S K^{\pm}\pi^{\mp}$.
- $B^0 \to K^0_S K^{\pm}\pi^{\mp}$.
- $B^+ \to D^0 h$ modes combined into one PDF.
- Combinatorial background.

Each 2D PDF component is the product of two PDFs, modelling the $m_{K^0_S\pi}$ and $m_{K^0_S K\pi}$ variables. The PDF shapes used are informed by the simulated event shapes shown in Chapter 4 and expectations based on the resonances in the decay. For example, backgrounds containing a real K^* use a relativistic Breit-Wigner shape for the $m_{K^0_S\pi}$ distribution, although with differences in the widths and means due to the constraint of the mass hypothesis to the $B^0_s \to K^0_S K^{\pm}\pi^{\mp}$ decay. The fixed values of the PDF parameters are determined from fits to simulated events, or side-band events in the case of the combinatorial background. The fits of these PDFs to the data samples are shown in Sections 4.2-4.5.

Many of the background modes have shapes which are not easily parameterised. For these shapes a RooKeysPdf was created using the simulated events for the mode. This PDF is generated via a superposition of Gaussian kernels and can quickly fit non-parameterisable shapes [68]. A summary of the PDF shapes used for the background modes is given in Table 5.2.

From fits to toy experiments it was found that allowing all of these background PDF yields to float freely caused the fit to be unstable. Analysis of the pull distributions and parameter correlation coefficients revealed that the $B^0_s \to K^*(K\pi^0)K^*(K^0_S\pi)$ and $B \to K^*\pi\pi$ PDF yields were the cause. To fix the instability, these yields are parameterised by the $B^0 \to K^{*+}\pi^-$ PDF yield. The relative yields

Table 5.2: Summary of PDF shapes used for the $B^0_{d,s} \to K^{*\pm}K^{\mp}$ background components.

Background	$m_{\rm B}$ PDF	$m_{K^0_S\pi}$ PDF
$B^0 \to K^{*+}\pi^-$	Crystal Ball	Relativistic Breit-Wigner
$\Lambda^0_b \to K^{*-}p$	RooKeysPdf	Relativistic Breit-Wigner
$B \to K^*\pi\pi$	RooKeysPdf	Relativistic Breit-Wigner
$B^0_s \to K^*(K\pi^0)K^*(K^0_S\pi)$	RooKeysPdf	Relativistic Breit-Wigner
$B^0_s \to K^0_S K^{\pm}\pi^{\mp}$	Crystal Ball	Linear
$B^0 \to K^0_S K^{\pm}\pi^{\mp}$	Crystal Ball	Linear
$B^+ \to D^0 h$	RooKeysPdf	Linear
Combinatorial	Linear	Linear

calculated in Section 4.6 are used to specify the yields relative to the $B^0 \to K^{*+}\pi^-$ mode.

Also, the non-resonant yields are constrained in several ways. The non-resonant Crystal Ball shapes share the floating mean and width parameters of the signal $B^0_{d,s} \to K^{*\pm}K^{\mp}$ modes. However, they do have different fixed values for the tail parameters.

Also, the $B^0_s \to K^0_S K^{\pm}\pi^{\mp}$ yields are forced to have the same ratio to the $B^0_s \to K^{*\pm}K^{\mp}$ signal yields between the Long-Long and Down-Down K^0_S samples. The same constraint was applied to the $B^0 \to K^0_S K^{\pm}\pi^{\mp}$ and $B^0 \to K^{*\pm}K^{\mp}$ yields. There is no physical reason why these ratios should differ between the K^0_S samples, so this constraint should only prevent the fit from using a minimum which is unphysical.

The linear shapes used in this analysis are not simple linear functions passing through the origin. A linear PDF should never be allowed to give a negative value as this is unphysical and causes unpredictable behaviour in the fitting algorithms. For the linear shapes with positive gradients in this analysis, the cases

$$P(x) = \begin{cases} m(x - C) & \text{if } x \geq C \\ 0 & \text{if } x < C \end{cases} \tag{5.7}$$

are used to parameterise the value of the PDF at a given value of the observable, x. Here, C is used as a threshold value for when the linear 'turns on' and m is the gradient. However, the gradient is cancelled during the normalisation conditions of the PDF. For linear shapes with negative gradients, such as the combinatorial background in the $m_{K^0_S K\pi}$ variable, the same parameterisation is used but with reversed inequalities.

The combinatorial background PDFs use two of these linear PDF shapes. An increasing one for the $m_{\mathrm{K^0_S}\pi}$ PDF and a decreasing one for the $m_{\mathrm{K^0_S}\mathrm{K}\pi}$ PDF. In the final fit it was found that the $m_{\mathrm{K^0_S}\mathrm{K}\pi}$ threshold value needed to be fixed, while the $m_{\mathrm{K^0_S}\pi}$ threshold value could be allowed to float. As shown in Fig. 4.16, the amount of combinatorial events in the upper side-bands was extremely small. A systematic is assigned to the final choice of the threshold value, along with the choices for the other fixed parameters, in Section 7.4.

The free parameters in the final fit model for the $\mathrm{K^0_S K^\pm \pi^\mp}$ final state are given in Table 5.3. Any parameter without a $\mathrm{K^0_S}$ reconstruction category label (DD or LL) is shared between the models during the simultaneous fit.

Table 5.3: Free fit model parameters for the $\mathrm{K^0_S K^\pm \pi^\mp}$ final state.

Free parameter symbol	Description
$N^{DD}_{\mathrm{B^0_s}\to\mathrm{K^{*\pm}K^\mp}}$	Yield of Down-Down $\mathrm{B^0_s}\to\mathrm{K^{*\pm}K^\mp}$.
$N^{LL}_{\mathrm{B^0_s}\to\mathrm{K^{*\pm}K^\mp}}$	Yield of Long-Long $\mathrm{B^0_s}\to\mathrm{K^{*\pm}K^\mp}$.
$N^{DD}_{\mathrm{B^0}\to\mathrm{K^{*\pm}K^\mp}}$	Yield of Down-Down $\mathrm{B^0}\to\mathrm{K^{*\pm}K^\mp}$.
$N^{LL}_{\mathrm{B^0}\to\mathrm{K^{*\pm}K^\mp}}$	Yield of Long-Long $\mathrm{B^0}\to\mathrm{K^{*\pm}K^\mp}$.
$N^{DD}_{\mathrm{B^0}\to\mathrm{K^{*+}}\pi^-}$	Yield of Down-Down $\mathrm{B^0}\to\mathrm{K^{*+}}\pi^-$.
$N^{LL}_{\mathrm{B^0}\to\mathrm{K^{*+}}\pi^-}$	Yield of Long-Long $\mathrm{B^0}\to\mathrm{K^{*+}}\pi^-$.
$N^{DD}_{\Lambda^0_b\to\mathrm{K^{*-}p}}$	Yield of Down-Down $\Lambda^0_b\to\mathrm{K^{*-}p}$.
$N^{LL}_{\Lambda^0_b\to\mathrm{K^{*-}p}}$	Yield of Long-Long $\Lambda^0_b\to\mathrm{K^{*-}p}$.
$N^{DD}_{\mathrm{B^+}\to\mathrm{D^0h}}$	Yield of Down-Down $\mathrm{B^+}\to\mathrm{D^0h}$.
$N^{LL}_{\mathrm{B^+}\to\mathrm{D^0h}}$	Yield of Long-Long $\mathrm{B^+}\to\mathrm{D^0h}$.
N^{DD}_{Comb}	Yield of Down-Down Combinatorial.
N^{LL}_{Comb}	Yield of Long-Long Combinatorial.
$R_{\mathrm{B^0}\to\mathrm{K^0_S K^\pm}\pi^\mp}$	Ratio of the $\mathrm{B^0}\to\mathrm{K^{*\pm}K^\mp}$ to $\mathrm{B^0}\to\mathrm{K^0_S K^\pm}\pi^\mp$ yields.
$R_{\mathrm{B^0_s}\to\mathrm{K^0_S K^\pm}\pi^\mp}$	Ratio of the $\mathrm{B^0_s}\to\mathrm{K^{*\pm}K^\mp}$ to $\mathrm{B^0_s}\to\mathrm{K^0_S K^\pm}\pi^\mp$ yields.
$x^{\mathrm{DD}}_{\mathrm{K^*}}$	Threshold value of the linear $m_{\mathrm{K^0_S}\pi}$ PDF for Down-Down Combinatorial.
$x^{\mathrm{LL}}_{\mathrm{K^*}}$	Threshold value of the linear $m_{\mathrm{K^0_S}\pi}$ PDF for Long-Long Combinatorial.
$\bar{m}_{\mathrm{B^0_s}}$	Mean of the $\mathrm{B^0_s}\to\mathrm{K^{*\pm}K^\mp}$ $\mathrm{B^0_s}$ mass peak.
$\bar{m}_{\mathrm{K^*}}$	Mean of the $\mathrm{B^0_s}\to\mathrm{K^{*\pm}K^\mp}$ $\mathrm{K^*}$ mass peak.
$\sigma_{\mathrm{B^0_s}}$	Width of the $\mathrm{B^0_s}\to\mathrm{K^{*\pm}K^\mp}$ $\mathrm{B^0_s}$ mass peak.
$\sigma_{\mathrm{K^*}}$	Width of the $\mathrm{B^0_s}\to\mathrm{K^{*\pm}K^\mp}$ $\mathrm{K^*}$ mass peak.

5.5 Simulation studies

To validate the final fit model, pull distributions for the fitted parameters were calculated. The RooMCStudy class in the software package RooFit [69] was used. This class interfaces with RooFit and streamlines the generation and fitting of many independent toy samples. The resulting pull distributions for 1000 toy experiments are shown in Figs. 5.4-5.5. The mean and widths of the fitted Gaussian PDFs to the pull distributions are summarised in Table 5.4. These suggest that the fit model has only small bias on the signal yields, which is assigned a systematic uncertainty in Section 7.4.

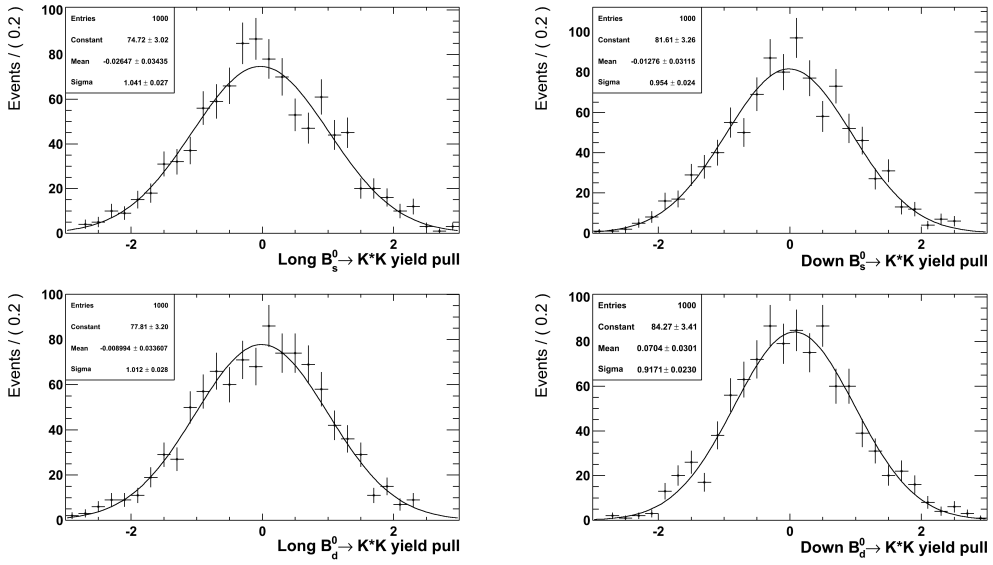

Figure 5.4: Pull distributions of the $B^0_{d,s} \to K^{*\pm}K^{\mp}$ yield parameters. Fitted with Gaussian PDFs.

Table 5.4: Pull distribution means and widths from toy experiments for the $B^0_{d,s} \to K^{*\pm}K^{\mp}$ yield parameters.

Yield parameter	Pull mean	Pull width
$B^0_s \to K^{*\pm}K^{\mp}$ (Long-Long K^0_S)	-0.03 ± 0.03	1.04 ± 0.03
$B^0_s \to K^{*\pm}K^{\mp}$ (Down-Down K^0_S)	-0.01 ± 0.03	0.95 ± 0.02
$B^0 \to K^{*\pm}K^{\mp}$ (Long-Long K^0_S)	-0.01 ± 0.03	1.01 ± 0.03
$B^0 \to K^{*\pm}K^{\mp}$ (Down-Down K^0_S)	0.07 ± 0.03	0.92 ± 0.02

Significant correlations between the fitted parameters are also a good indi-

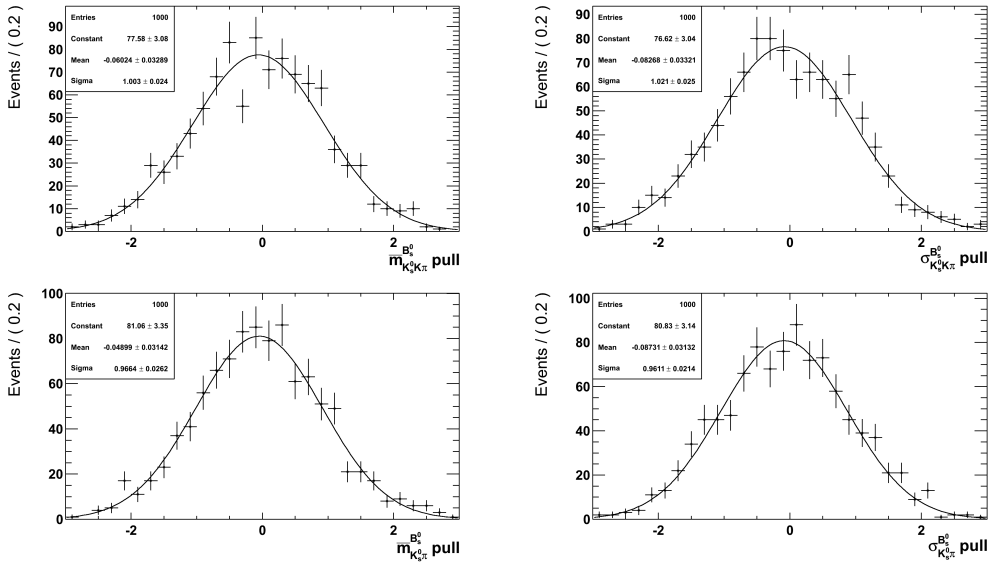

Figure 5.5: Pull distributions of the $B^0_{d,s} \to K^{*\pm}K^{\mp}$ mean and sigma shape parameters. Fitted with Gaussian PDFs.

cator of bias in the fit model. To evaluate the correlation coefficients of the fitted parameters, the fitted values from the pull calculation were used. For each pair of parameters, the fitted values were placed into a 2D histogram. The correlation coefficient was then evaluated from this. The resulting factors are displayed in Tables 5.5-5.7, where the correlations between the Long-Long K^0_S and Down-Down K^0_S yields have been omitted for brevity as they are all, as expected, small. The ratio parameters $R_{B^0_s \to K^0_S K^{\pm}\pi^{\mp}}$ and $R_{B^0 \to K^0_S K^{\pm}\pi^{\mp}}$ affect both the Long-Long and Down-Down signal yields, they are therefore included in both tables.

The only source of correlation with an absolute value larger than 0.5 comes from the Long-Long combinatorial yield, N^{LL}_{Comb} and the Long-Long $B^+ \to D^0h$ yield, $N^{LL}_{B^+ \to D^0 h}$. This has a negative correlation of -0.71 suggesting that the combinatorial background and $B^+ \to D^0h$ PDFs are being varied by the fit against one another. This was assumed to be due to the approximation of the combinatorial shape as a linear function in the $m_{K^0_S K\pi}$ variable and the lack of information about the combinatorial PDF shape at the lower edge of the fit range. Given that the $B^+ \to D^0h$ shape in the $m_{K^0_S K\pi}$ spectrum is far from the important B^0_s mass peak, it was felt that this effect could be safely neglected.

Table 5.5: Correlation coefficients for the Long-Long K^0_S yield parameters.

	$N^{LL}_{B^0_s\to K^*K}$	$N^{LL}_{B^0\to K^*K}$	$R_{B^0\to K^0_S K^\pm\pi^\mp}$	$R_{B^0_s\to K^0_S K^\pm\pi^\mp}$	$N^{LL}_{B^0\to K^*\pi}$	$N^{LL}_{B^+\to D^0 h}$	$N^{LL}_{\Lambda^0_b\to K^{*-}p}$	N^{LL}_{Comb}
$N^{LL}_{B^0_s\to K^*K}$	-	0.14	-0.03	-0.07	-0.13	0.23	-0.19	-0.11
$N^{LL}_{B^0\to K^*K}$	0.14	-	0.04	-0.07	-0.18	0.08	-0.13	-0.23
$R_{B^0\to K^0_S K^\pm\pi^\mp}$	-0.03	0.04	-	-0.01	-0.03	0.01	-0.01	-0.03
$R_{B^0_s\to K^0_S K^\pm\pi^\mp}$	-0.07	-0.07	-0.01	-	0.05	-0.06	0.05	0.07
$N^{LL}_{B^0\to K^*\pi}$	-0.13	-0.18	-0.03	0.05	-	-0.29	0.21	0.40
$N^{LL}_{B^+\to D^0 h}$	0.23	0.08	0.01	-0.06	-0.29	-	-0.10	-0.71
$N^{LL}_{\Lambda^0_b\to K^{*-}p}$	-0.19	-0.13	-0.01	0.05	0.21	-0.10	-	0.12
N^{LL}_{Comb}	-0.11	-0.23	-0.03	0.07	0.40	-0.71	0.12	-

Table 5.6: Correlation coefficients for the Down-Down K^0_S yield parameters.

	$N^{DD}_{B^0_s\to K^*K}$	$N^{DD}_{B^0\to K^*K}$	$R_{B^0\to K^0_S K^\pm\pi^\mp}$	$R_{B^0_s\to K^0_S K^\pm\pi^\mp}$	$N^{DD}_{B^0\to K^*\pi}$	$N^{DD}_{B^+\to D^0 h}$	$N^{DD}_{\Lambda^0_b\to K^{*-}p}$	N^{DD}_{Comb}
$N^{DD}_{B^0_s\to K^*K}$	-	0.19	-0.06	-0.02	0.13	-0.15	-0.27	0.20
$N^{DD}_{B^0\to K^*K}$	0.19	-	-0.04	-0.24	0.14	0.08	-0.10	-0.08
$R_{B^0\to K^0_S K^\pm\pi^\mp}$	-0.06	-0.04	-	-0.01	-0.06	0.05	0.07	-0.05
$R_{B^0_s\to K^0_S K^\pm\pi^\mp}$	-0.02	-0.24	-0.01	-	0.00	-0.15	0.02	0.21
$N^{DD}_{B^0\to K^*\pi}$	0.13	0.14	-0.06	0.00	-	-0.12	-0.29	0.14
$N^{DD}_{B^+\to D^0 h}$	-0.15	0.08	0.05	-0.15	-0.12	-	0.22	-0.19
$N^{DD}_{\Lambda^0_b\to K^{*-}p}$	-0.27	-0.10	0.07	0.02	-0.29	0.22	-	-0.28
N^{DD}_{Comb}	0.20	-0.08	-0.05	0.21	0.14	-0.19	-0.28	-

Table 5.7: Correlation coefficients for the shape parameters.

	$\bar{m}_{\mathrm{B}^0_s}$	$\bar{m}_{\mathrm{K}^*}$	$\sigma_{\mathrm{B}^0_s}$	σ_{K^*}	$x^{\mathrm{LL}}_{\mathrm{K}^*}$	$x^{\mathrm{DD}}_{\mathrm{K}^*}$
$\bar{m}_{\mathrm{B}^0_s}$	-	-0.03	0.25	0.21	-0.22	-0.06
$\bar{m}_{\mathrm{K}^*}$	-0.03	-	-0.03	0.12	0.011	0.01
$\sigma_{\mathrm{B}^0_s}$	0.25	-0.03	-	0.29	-0.21	-0.08
σ_{K^*}	0.21	0.12	0.29	-	-0.01	0.01
$x^{\mathrm{LL}}_{\mathrm{K}^*}$	-0.22	0.01	-0.21	-0.01	-	0.01
$x^{\mathrm{DD}}_{\mathrm{K}^*}$	-0.06	0.01	-0.08	0.01	0.01	-

5.6 $\mathrm{B}^0_{d,s} \to \mathrm{K}^{*\pm}\pi^{\mp}$ model

The fit model for the $\mathrm{K}^0_{\mathrm{S}}\pi^+\pi^-$ final state is extremely similar to the $\mathrm{K}^0_{\mathrm{S}}\mathrm{K}^{\pm}\pi^{\mp}$ fit model. Therefore the same level of detail is not included in this description, while any differences are highlighted.

Signal PDFs

A Crystal Ball and Breit-Wigner PDF shape are again used to parameterise the signal $\mathrm{B}^0_{d,s} \to \mathrm{K}^{*\pm}\pi^{\mp}$ modes. The $\mathrm{B}^0 \to \mathrm{K}^{*+}\pi^-$ mode is expected to be dominant compared to the $\mathrm{B}^0_s \to \mathrm{K}^{*-}\pi^+$. Therefore the same parameterisation of the $\mathrm{B}^0_s \to \mathrm{K}^{*-}\pi^+$ means and widths in terms of the corresponding $\mathrm{B}^0 \to \mathrm{K}^{*+}\pi^-$ parameters is made. All of the fixed shape parameter values are taken from the simulated events. Gaussian constraints, informed by the fits to the simulated events, are applied to the free shape parameters of the signal PDFs.

Background PDFs

The background modes considered for this final state are:

- $\mathrm{B}^0_s \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$.
- $\mathrm{B}^0 \to \mathrm{K}^*(\mathrm{K}^0_{\mathrm{S}}\pi^+)\rho(\pi^-\pi^0)$ and $\mathrm{B}^+ \to \mathrm{K}^*(\mathrm{K}^0_{\mathrm{S}}\pi^+)\pi^+\pi^-$ combined into one PDF, abbreviated as $\mathrm{B} \to \mathrm{K}^*\pi\pi$
- $\mathrm{B}^0_s \to \mathrm{K}^0_{\mathrm{S}}\pi^+\pi^-$.
- $\mathrm{B}^0 \to \mathrm{K}^0_{\mathrm{S}}\pi^+\pi^-$.
- $\mathrm{B}^+ \to \mathrm{D}^0\pi^+$.

- Combinatorial background.

All of these background PDF shapes are the same as in the $K^0_S K^\pm \pi^\mp$ final state, which are listed in Table 5.2. For this model, the smaller number of PDFs allowed all of the component PDF yields to float freely. However, the non-resonant $B^0_s \to K^0_S \pi^+ \pi^-$ yield is still forced to have the same ratio to the $B^0_s \to K^{*-} \pi^+$ yield between the Long-Long and Down-Down K^0_S samples. As before, it also shared the $B^0_s \to K^{*-} \pi^+$ Crystal Ball mean and width parameters. These constraints are also applied to the $B^0 \to K^0_S \pi^+ \pi^-$ and $B^0 \to K^{*+} \pi^-$ PDFs.

5.7 Fit results on data

After the fit models to the $K^0_S K^\pm \pi^\mp$ and $K^0_S \pi^+ \pi^-$ final states were created and validated, they were used to fit to the collision data samples. The results of each fit are given in Tables 5.8-5.9. Here there are non-zero signal yields for the unmeasured $B^0_s \to K^{*\pm} K^\mp$, $B^0_s \to K^{*-} \pi^+$ signal modes and for the previously measured $B^0 \to K^{*+} \pi^-$ normalisation mode. The suppressed $B^0 \to K^{*\pm} K^\mp$ mode yields are consistent with zero. The statistical and total significances of these signal yields are calculated in Chapter 8, along with the relative branching fractions and limits.

The fitted PDF components and data samples are plotted in Figs. 5.6-5.7 for both the $m_{K^0_S K \pi}$ ($m_{K^0_S \pi\pi}$) and $m_{K^0_S \pi}$ projections. Here the favoured signal mode in both the $K^0_S K^\pm \pi^\mp$ and $K^0_S \pi^+ \pi^-$ final states is always shown in red with the non-resonant component in maroon. The suppressed mode is shown in pink, with the non-resonant component in purple. The cross-feed from decays with a misidentified bachelor track is shown in cyan. The green PDF component is the combinatorial background and orange is the combined background components for the remaining PDFs.

Table 5.8: Fitted parameter values after the fit to data for the $\mathrm{K^0_S K^\pm \pi^\mp}$ final state. Only statistical uncertainties from the fit are given.

Fit parameter	Fitted value
$N^{DD}_{\mathrm{B^0_s \to K^{*\pm} K^\mp}}$	62 ± 10
$N^{LL}_{\mathrm{B^0_s \to K^{*\pm} K^\mp}}$	40 ± 8
$N^{DD}_{\mathrm{B^0 \to K^{*\pm} K^\mp}}$	4 ± 3
$N^{LL}_{\mathrm{B^0 \to K^{*\pm} K^\mp}}$	0 ± 4
$N^{DD}_{\mathrm{B^0 \to K^{*+} \pi^-}}$	0.8 ± 0.8
$N^{LL}_{\mathrm{B^0 \to K^{*+} \pi^-}}$	2.8 ± 3.1
$N^{DD}_{\mathrm{\Lambda_b^0 \to K^{*-} p}}$	-2.2 ± 2.5
$N^{LL}_{\mathrm{\Lambda_b^0 \to K^{*-} p}}$	6 ± 7
$N^{DD}_{\mathrm{B^+ \to D^0 h}}$	-12 ± 8
$N^{LL}_{\mathrm{B^+ \to D^0 h}}$	-11 ± 7
N^{DD}_{Comb}	99 ± 17
N^{LL}_{Comb}	231 ± 22
$R_{\mathrm{B^0 \to K^0_S K^\pm \pi^\mp}}$	0.0 ± 1.0
$R_{\mathrm{B^0_s \to K^0_S K^\pm \pi^\mp}}$	0.41 ± 0.16
$x^{\mathrm{DD}}_{\mathrm{K^*}}$	636 ± 50
$x^{\mathrm{LL}}_{\mathrm{K^*}}$	-400 ± 800
$\bar{m}_{\mathrm{B^0_s}}$	5369.9 ± 1.6
$\bar{m}_{\mathrm{K^*}}$	891.7 ± 2.2
$\sigma_{\mathrm{B^0_s}}$	18.5 ± 1.3
$\sigma_{\mathrm{K^*}}$	49.1 ± 2.4

Table 5.9: Fitted parameter values after the fit to data for the $\mathrm{K}^0_\mathrm{S}\pi^+\pi^-$ final state. Only statistical uncertainties from the fit are given.

Fit parameter	Fitted value
$N^{DD}_{\mathrm{B}^0\to\mathrm{K}^{*+}\pi^-}$	165 ± 16
$N^{LL}_{\mathrm{B}^0\to\mathrm{K}^{*+}\pi^-}$	80 ± 10
$N^{DD}_{\mathrm{B}^0_s\to\mathrm{K}^{*-}\pi^+}$	23 ± 8
$N^{LL}_{\mathrm{B}^0_s\to\mathrm{K}^{*-}\pi^+}$	5 ± 4
$N^{DD}_{\mathrm{B}^0_s\to\mathrm{K}^{*\pm}\mathrm{K}^\mp}$	-4 ± 6
$N^{LL}_{\mathrm{B}^0_s\to\mathrm{K}^{*\pm}\mathrm{K}^\mp}$	3 ± 7
$N^{DD}_{\mathrm{B}^+\to\mathrm{D}^0\pi^+}$	52 ± 16
$N^{LL}_{\mathrm{B}^+\to\mathrm{D}^0\pi^+}$	22 ± 13
$N^{DD}_{\mathrm{B}^0\to\mathrm{K}^{*+}\rho^-}$	26 ± 10
$N^{LL}_{\mathrm{B}^0\to\mathrm{K}^{*+}\rho^-}$	38 ± 11
N^{DD}_{Comb}	190 ± 27
N^{LL}_{Comb}	276 ± 26
$R_{\mathrm{B}^0\to\mathrm{K}^0_\mathrm{S}\pi^+\pi^-}$	0.79 ± 0.14
$R_{\mathrm{B}^0_s\to\mathrm{K}^0_\mathrm{S}\pi^+\pi^-}$	0.6 ± 0.4
$x^{\mathrm{DD}}_{\mathrm{K}^*}$	360 ± 210
$x^{\mathrm{LL}}_{\mathrm{K}^*}$	-1400 ± 1800
$\bar{m}_{\mathrm{B}^0}$	5282.4 ± 1.0
$\bar{m}_{\mathrm{K}^*}$	891.1 ± 2
σ_{B^0}	18.5 ± 1.0
σ_{K^*}	49.1 ± 2.3

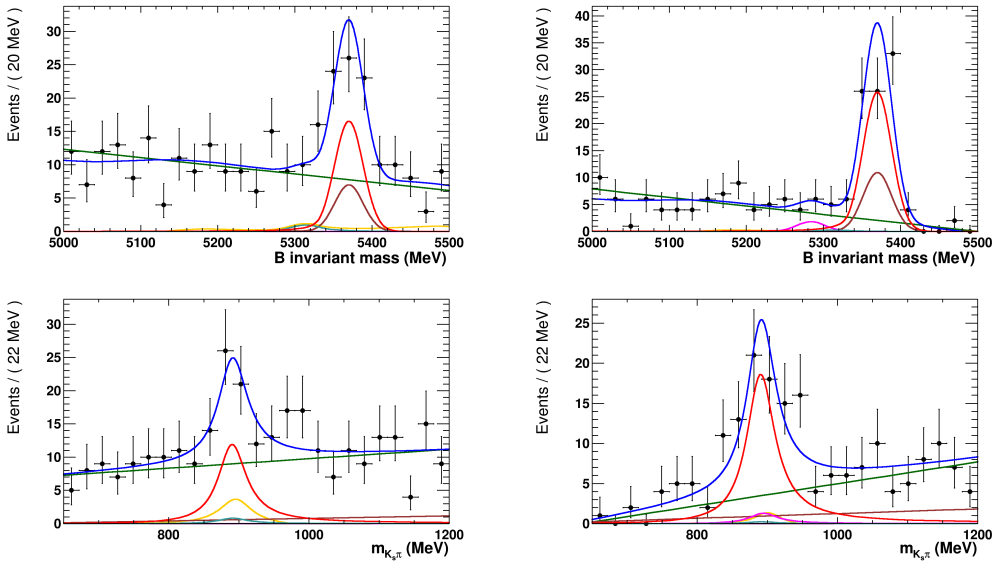

Figure 5.6: Plots of the $K^0_S K^{\pm}\pi^{\mp}$ final state data sample with the fitted PDFs projected into each observable. The total PDF projection for each K^0_S category is shown in blue. The signal decay PDFs, $B^0_s \to K^{*\pm}K^{\mp}$ decay (Red) and $B^0 \to K^{*\pm}K^{\mp}$ (Pink) are shown. The non-resonant components $B^0_s \to K^0_S K^{\pm}\pi^{\mp}$ (Maroon) and $B^0 \to K^0_S K^{\pm}\pi^{\mp}$ (Purple) are plotted. The $B^0 \to K^{*+}\pi^-$ cross-feed is shown in Cyan. The green PDF component is the combinatorial background and orange is the combined background components for the remaining PDFs. Top: $m_{K^0_S K\pi}$ spectrum. Bottom: $m_{K^0_S \pi}$ spectrum. Left: Long-Long K^0_S sample. Right: Down-Down K^0_S sample.

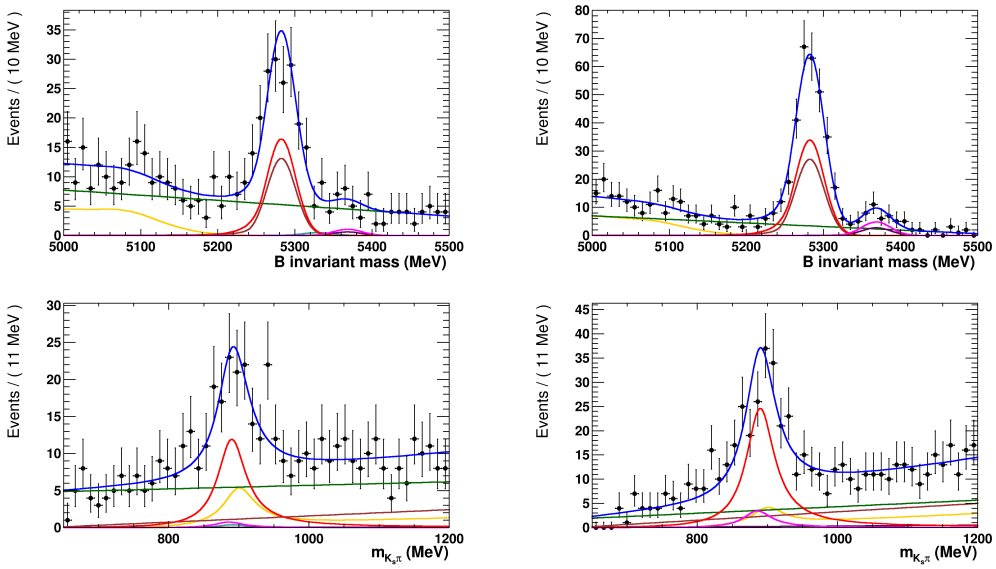

Figure 5.7: Plots of the $K^0_S\pi^+\pi^-$ final state data sample with the fitted PDFs projected into each observable. The total PDF projection for each K^0_S category is shown in blue. The signal decay PDFs, $B^0 \to K^{*+}\pi^-$ decay (Red) and $B^0_s \to K^{*-}\pi^+$ (Pink) are shown. The non-resonant components $B^0 \to K^0_S\pi^+\pi^-$ (Maroon) and $B^0_s \to K^0_S\pi^+\pi^-$ (Purple) are plotted. The $B^0_s \to K^{*\pm}K^{\mp}$ cross-feed is shown in Cyan. The green PDF component is the combinatorial background and orange is the combined background components for the remaining PDFs. Top: $m_{K^0_S\pi\pi}$ spectrum. Bottom: $m_{K^0_S\pi}$ spectrum. Left: Long-Long K^0_S sample. Right: Down-Down K^0_S sample.

6 Efficiencies

6.1 Introduction

Using the fit models described in Chapter 5 to extract the yields of various modes is only one part of the measurement of the branching fractions. Obviously not all of the decays produced at the LHCb interaction point are in the final data sample. The decay products might not be within the LHCb acceptance, the tracks might not be reconstructed, or they may not be identified as the correct particle.

These issues cause the data sample selection to have several sources of inefficiency. Knowledge of each of the contributing efficiencies to each yield is necessary to calculate the branching fractions. These are detailed in the master formula and determined during this chapter.

Master formula of branching ratio

The master formula for the extraction of the ratio of branching fractions is

$$\frac{\mathcal{B}\left(\mathrm{B}^0_{d,s}\to \mathrm{K}^{*\pm}\mathrm{h}^{\mp}\right)}{\mathcal{B}\left(\mathrm{B}^0\to \mathrm{K}^{*+}\pi^-\right)} = \frac{f_d}{f_{d,s}}\cdot\frac{\varepsilon^{\mathrm{Acc}}_{\mathrm{B}^0\to \mathrm{K}^{*+}\pi^-}}{\varepsilon^{\mathrm{Acc}}_{\mathrm{B}^0_{d,s}\to \mathrm{K}^{*\pm}\mathrm{h}^{\mp}}}\cdot\frac{\varepsilon^{\mathrm{Sel}}_{\mathrm{B}^0\to \mathrm{K}^{*+}\pi^-}}{\varepsilon^{\mathrm{Sel}}_{\mathrm{B}^0_{d,s}\to \mathrm{K}^{*\pm}\mathrm{h}^{\mp}}}\cdot\frac{\varepsilon^{\mathrm{PID}}_{\mathrm{B}^0\to \mathrm{K}^{*+}\pi^-}}{\varepsilon^{\mathrm{PID}}_{\mathrm{B}^0_{d,s}\to \mathrm{K}^{*\pm}\mathrm{h}^{\mp}}}\cdot\frac{N_{\mathrm{B}^0_{d,s}\to \mathrm{K}^{*\pm}\mathrm{h}^{\mp}}}{N_{\mathrm{B}^0\to \mathrm{K}^{*+}\pi^-}}, \tag{6.1}$$

where $\varepsilon^{\rm Acc}$ is the LHCb acceptance efficiency, $\varepsilon^{\rm PID}$ is the PID efficiency and $\varepsilon^{\rm Sel}$ is the combined trigger, reconstruction, stripping and offline selection efficiencies.

This formula follows from Eq. 4.1 where the branching fraction of a signal mode is taken relative to the $\rm B^0 \to K^{*+}\pi^-$ normalisation mode. Measuring the three absolute efficiencies for both the signal and normalisation modes allows the ratios $\epsilon_{\rm B^0\to K^{*+}\pi^-}/\epsilon_{\rm B^0_{d,s}\to K^{*\pm}h^\mp}$ to be calculated, which can then be used in the master formula. The measurements of the three ratios are described in Sections 6.2-6.4.

6.2 Detector acceptance

As discussed in Section 2.3, the LHCb detector has a pseudorapidity acceptance of $2 \lesssim \eta \lesssim 5$ in the positive z direction. However, the B mesons, and their decay products, will not necessarily be within this acceptance range. The tracks of the daughter particles allow the reconstruction of the B decay. If these tracks fall outside the acceptance then the reconstruction is never applied to them.

During the production of the simulated $\rm B^0_{d,s} \to K^{*\pm}K^\mp$ events the decay products which pass the acceptance requirements are monitored, providing a measurement of the efficiency. In order to evaluate the acceptance efficiency of the $\rm B^0_{d,s} \to K^{*\pm}\pi^\mp$ modes, the LHCb simulation package Gauss was used. Ten thousand $\rm B^0_{d,s} \to K^{*\pm}\pi^\mp$ decays were generated for each mode and each of the magnet polarities. The efficiency of applying the acceptance requirements to decay products was calculated from the number of passing and failing decays. The absolute efficiencies for the $\rm B^0_{d,s} \to K^{*\pm}h^\mp$ modes are given in Table 6.1, while the efficiency ratios relative to the $\rm B^0 \to K^{*+}\pi^-$ mode are given in Table 6.2. In both tables, only statistical uncertainties are given.

Table 6.1: Detector acceptance efficiencies of the $\rm B^0_{d,s} \to K^{*\pm}h^\mp$ modes. Calculated from simulated events.

Decay mode	MagDown efficiency (%)	MagUp efficiency (%)
$\rm B^0 \to K^{*\pm}K^\mp$	18.59 ± 0.06	18.70 ± 0.06
$\rm B^0_s \to K^{*\pm}K^\mp$	18.62 ± 0.06	18.68 ± 0.06
$\rm B^0 \to K^{*+}\pi^-$	18.93 ± 0.24	18.15 ± 0.23
$\rm B^0_s \to K^{*-}\pi^+$	18.34 ± 0.23	18.60 ± 0.24

Table 6.2: Detector acceptance efficiencies for the $B^0_{d,s} \to K^{*\pm}K^{\mp}$ modes, relative to the $B^0 \to K^{*+}\pi^-$ mode.

Decay mode	MagDown relative efficiency	MagUp relative efficiency
$B^0 \to K^{*\pm}K^{\mp}$	1.02 ± 0.01	0.97 ± 0.01
$B^0_s \to K^{*\pm}K^{\mp}$	1.02 ± 0.01	0.97 ± 0.01
$B^0_s \to K^{*-}\pi^+$	1.03 ± 0.02	0.98 ± 0.02

6.3 Reconstruction and selection

Both the reconstruction of the events and the subsequent selection requirements made by the stripping and offline selections have associated efficiencies. Prior to the PID requirements all selection cuts are the same between the decay modes. This allowed the reconstruction and selection efficiencies to be evaluated as a combined efficiency, with the PID efficiencies calculated separately in Section 6.4. However, due to the different reconstruction efficiencies and selection requirements of the Down-Down K^0_S and Long-Long K^0_S candidates, these efficiencies are separately determined.

The selection efficiencies are calculated by applying the reconstruction and selection requirements to simulated events that have passed the acceptance requirements. The selection is directly applied to $B^0_{d,s} \to K^0_S\pi^+\pi^-$ events that are reweighted to resemble $B^0_{d,s} \to K^{*\pm}\pi^{\mp}$ decays. The weights are calculated from the ratios of amplitude squares, using the LAURA++ package. Hereafter these events are simply referred to as the $B^0_{d,s} \to K^{*\pm}\pi^{\mp}$ samples. In the calculation of the efficiencies, the ROOT `TEfficiency` class [70] is used with the confidence intervals calculated with the Clopper-Pearson method [71]. The absolute combined reconstruction and selection efficiencies are given in Table 6.3 and the efficiencies relative to the $B^0 \to K^{*+}\pi^-$ mode are given in Table 6.4. Only the statistical uncertainties are quoted here.

6.4 PID efficiencies

As mentioned in Section 3.7 it is not possible to find the correct PID efficiency by naively applying the DLL selections, described in Table 3.12, to the simulated events. The DLL values produced by the RICH detectors are highly sensitive to the number of tracks in the event, as well as the resulting emission and detection of the Cherenkov light. The effect of small discrepancies in the simulated events is

Table 6.3: Absolute reconstruction and selection efficiencies for the $B^0_{d,s} \to K^{*\pm}K^{\mp}$ and $B^0_{d,s} \to K^{*\pm}\pi^{\mp}$ modes.

Decay mode	MagUp efficiency (%)	MagDown efficiency (%)
$B^0_s \to K^{*\pm}K^{\mp}$ Long-Long	0.152 ± 0.004	0.150 ± 0.004
$B^0 \to K^{*\pm}K^{\mp}$ Long-Long	0.139 ± 0.004	0.149 ± 0.004
$B^0_s \to K^{*-}\ \pi^+$ Long-Long	0.138 ± 0.007	0.147 ± 0.007
$B^0 \to K^{*+}\pi^-$ Long-Long	0.122 ± 0.006	0.142 ± 0.007
$B^0_s \to K^{*\pm}K^{\mp}$ Down-Down	0.292 ± 0.005	0.294 ± 0.005
$B^0 \to K^{*\pm}K^{\mp}$ Down-Down	0.286 ± 0.005	0.293 ± 0.005
$B^0_s \to K^{*-}\ \pi^+$ Down-Down	0.263 ± 0.009	0.263 ± 0.009
$B^0 \to K^{*+}\pi^-$ Down-Down	0.257 ± 0.009	0.254 ± 0.009

Table 6.4: Combined reconstruction and selection efficiencies for the $B^0_s \to K^{*-}\ \pi^+$ and $B^0_{d,s} \to K^{*\pm}K^{\mp}$ modes, relative to the $B^0 \to K^{*+}\pi^-$ mode.

Decay mode	MagUp relative efficiency	MagDown relative efficiency
$B^0_s \to K^{*\pm}K^{\mp}$ Long-Long	0.81 ± 0.05	0.95 ± 0.06
$B^0 \to K^{*\pm}K^{\mp}$ Long-Long	0.88 ± 0.05	0.96 ± 0.06
$B^0_s \to K^{*-}\ \pi^+$ Long-Long	0.89 ± 0.07	0.97 ± 0.07
$B^0_s \to K^{*\pm}K^{\mp}$ Down-Down	0.88 ± 0.04	0.87 ± 0.04
$B^0 \to K^{*\pm}K^{\mp}$ Down-Down	0.90 ± 0.04	0.87 ± 0.04
$B^0_s \to K^{*-}\ \pi^+$ Down-Down	0.98 ± 0.05	0.97 ± 0.05

cumulative, causing the DLL distributions for the simulated and collision events to be significantly different. Therefore, a data driven method is used to evaluate the PID requirement efficiencies.

PID **efficiency method**

The `PIDCalib` package allows a user to re-weight calibration tracks of known ID (independent of the RICH DLL values) using variables to which the DLL distributions are sensitive e.g. momentum (p), transverse momentum (p_T), or number of tracks (nTracks). For pion and kaon tracks, these calibration tracks come from $D^{*+} \to D^0(K^-\pi^+)\pi^+$ decays. With these decays, pions and kaons can be identified without the need to use the RICH PID information. During the $B^0_{d,s} \to K^0_S h^{\pm} h'^{\mp}$ analysis the `PIDCalib` package could not produce average efficiencies for multiple track final states. [I] It could only give average efficiencies for each track separately and didn't take into account the possible correlations of the track kinematics. A tool was developed that can produce event-by-event efficiencies, which are then averaged to give the final PID efficiency for decay mode.

The `PIDCalib` tool allows the production of efficiency histograms for the track ID under analysis and the relevant DLL cuts. Efficiency histograms are binned in the p and p_T variables for the DLL cuts in Table 3.12. In this case, 16 bins in p and 4 bins in p_T are chosen, where the bin boundaries are manually tuned using $B^0 \to K^0_S\pi^+\pi^-$ simulated events to give approximately the same number of tracks in each bin. The histograms are made by binning the calibration tracks into a 2D histogram. Then the DLL cut is applied to them and a second histogram with the same binning is made with the resulting tracks. These two histograms are divided using the `TEfficiency` class, to calculate a 2D efficiency histogram.

The efficiency of the DLL cut may vary in time, due to the varying performance of the RICH sub-detectors over the course of the data taking. To account for this, different efficiency histograms are made with subsets of the calibration tracks, corresponding to the LHCb run ranges given in Table 6.5. In order to correctly account for this variation, the efficiency histograms for each run range should be weighted by the integrated luminosity of the run range they correspond to. As an approximation, a weighted average of the efficiency histograms is performed, using the formula

$$x_{\text{avg}} = \frac{w_1.x_1 + w_2.x_2}{w_1 + w_2}. \tag{6.2}$$

[I] This functionality is now possible in `PIDCalib` and the method follows the one used in this analysis.

Here the resulting average value of a bin, x_{avg}, from the same bin in two other histograms is calculated from their values, x_1 and x_2, and their weights w_1 and w_2. The weight for a bin is calculated from its uncertainty, σ, as

$$w = \frac{1}{\sigma^2}. \tag{6.3}$$

The integrated luminosity of the run range is correlated to the number of calibration tracks in that range. Therefore the statistical uncertainty of a bin is correlated to the integrated luminosity, making this a good approximation.

Table 6.5: Run ranges of the different `PIDCalib` Calibration tracks

MagUp Run Ranges	MagDown Run Ranges
89333 → 93516	87665 → 91992
93517 → 94234	91993 → 92735
94261 → 96641	92821 → 93224
96642 → 100243	93225 → 97884
100244 → 102505	97885 → 98198
102506 → 102893	98199 → 101905
102894 → 104263	101906 → 102378
-	102379 → 103361
-	103362 → 103686

With the efficiency histograms averaged, consider the PID efficiency measurement for $B_s^0 \to K^{*\pm}K^{\mp}$ events. A clean sample of $B_s^0 \to K^{*\pm}K^{\mp}$ events is required, so that the p and p_T distributions can be used to extract the PID efficiency. Since the p and p_T are well modelled in the simulated events, the larger sample of simulated events can be used instead of extracting the signal events from the collision data. This has the additional benefit that the simulated events will by definition be a clean sample. For each event, first the p and p_T for the kaon and pion tracks are found. These values are then used with the corresponding efficiency histograms to extract the two track efficiencies. These efficiencies are then multiplied together to form an event efficiency for every event in the sample. The efficiencies are then averaged to give the final value.

Statistical uncertainties

There are two separate sources of statistical uncertainty in the PID efficiency calculation. The first is the limited calibration track sample statistics which gives the uncertainties on the efficiency histogram bins. However, the dominant statistical

uncertainty comes from the limited number simulated events used in the calculation of the event efficiencies.

To estimate this source of uncertainty, 2D PDFs are found for the p vs. p_T distributions of each charged track, in each mode. These are created using the `RooNDKeysPdf` class and the simulated events for each mode. By sampling these PDFs, 500 independent toy data samples for each mode are created, each with the same number of events as used in the final PID efficiency calculation. By performing the PID efficiency calculation with each toy sample, a set of PID efficiencies is found for each mode. These sets are fitted with Gaussian PDFs and the widths are used as estimates of the statistical uncertainties on each measured PID efficiency. Two example histograms showing the fitted range of PID efficiencies are given in Fig. 6.1.

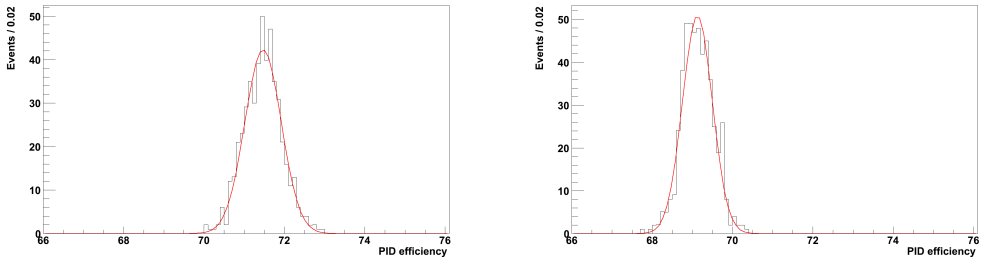

Figure 6.1: Fitted PID efficiencies for toy $B^0_s \to K^{*\pm}K^{\mp}$ magnet up events. Left: Long-Long K^0_S. Right: Down-Down K^0_S.

Measured efficiencies

Tables 6.6-6.7 give the measured values for the different $B^0_{d,s} \to K^{*\pm}h^{\mp}$ mode PID efficiencies, both absolute and relative to the $B^0 \to K^{*+}\pi^-$ mode. The quoted uncertainties are statistical only and are calculated by adding the two previously discussed uncertainties in quadrature.

Table 6.6: PID efficiencies for the $B^0_{d,s} \to K^{*\pm}K^{\mp}$ and $B^0_{d,s} \to K^{*\pm}\pi^{\mp}$ modes.

Decay mode	MagUp efficiency (%)	MagDown efficiency (%)
$B^0_s \to K^{*\pm}K^{\mp}$ Long-Long	71.7 ± 0.5	70.1 ± 0.5
$B^0 \to K^{*\pm}K^{\mp}$ Long-Long	71.7 ± 0.5	71.2 ± 0.5
$B^0_s \to K^{*-}\ \pi^+$ Long-Long	73.3 ± 0.9	73.1 ± 0.7
$B^0 \to K^{*+}\pi^-$ Long-Long	73.2 ± 0.7	74.1 ± 0.8
$B^0_s \to K^{*\pm}K^{\mp}$ Down-Down	69.3 ± 0.4	68.6 ± 0.3
$B^0 \to K^{*\pm}K^{\mp}$ Down-Down	69.9 ± 0.4	69.8 ± 0.4
$B^0_s \to K^{*-}\ \pi^+$ Down-Down	74.4 ± 0.5	74.6 ± 0.6
$B^0 \to K^{*+}\pi^-$ Down-Down	74.4 ± 0.6	74.7 ± 0.6

Table 6.7: PID efficiencies for the $B^0_{d,s} \to K^{*\pm}K^{\mp}$ and $B^0_s \to K^{*-}\ \pi^+$ modes, relative to the $B^0 \to K^{*+}\pi^-$ mode.

Decay mode	MagUp relative efficiency	MagDown relative efficiency
$B^0_s \to K^{*\pm}K^{\mp}$ Long-Long	1.02 ± 0.01	1.06 ± 0.01
$B^0 \to K^{*\pm}K^{\mp}$ Long-Long	1.02 ± 0.01	1.04 ± 0.01
$B^0_s \to K^{*-}\ \pi^+$ Long-Long	1.00 ± 0.02	1.01 ± 0.01
$B^0_s \to K^{*\pm}K^{\mp}$ Down-Down	1.07 ± 0.01	1.09 ± 0.01
$B^0 \to K^{*\pm}K^{\mp}$ Down-Down	1.06 ± 0.01	1.07 ± 0.01
$B^0_s \to K^{*-}\ \pi^+$ Down-Down	1.00 ± 0.01	1.00 ± 0.01

7 Systematics

7.1 Introduction

This chapter describes the different sources of systematic uncertainty affecting the various measurements in this analysis. The systematic uncertainties affecting the measurement of the yields and the relative branching fractions are summarised at the end of this chapter in Section 7.5.

7.2 Trigger requirements

The trigger has two potential sources of uncertainty for this analysis. The efficiency of the hardware hadron trigger on simulated events may be different when compared to the true trigger efficiency. Also, the ageing of the calorimeters during 2011 caused a relative shift of $\sim 10\%$ on the absolute trigger efficiencies. The systematic uncertainties on the relative branching fractions are here taken from the $\mathrm{B}^0_{d,s} \to \mathrm{K}^0_\mathrm{S}\mathrm{h}^\pm\mathrm{h}'^\mp$ analysis [1] and are listed in Table 7.1.

Table 7.1: Relative uncertainties on the ratio of branching fractions due to the trigger modelling uncertainty in the simulated events.

Relative B.F.	Long-Long K^0_S	Down-Down K^0_S
$\frac{\mathcal{B}(B^0_s \to K^{*\pm} K^{\mp})}{\mathcal{B}(B^0 \to K^{*+} \pi^-)}$	1.7%	1.4%
$\frac{\mathcal{B}(B^0 \to K^{*\pm} K^{\mp})}{\mathcal{B}(B^0 \to K^{*+} \pi^-)}$	1.1%	3.4%
$\frac{\mathcal{B}(B^0_s \to K^{*-} \pi^+)}{\mathcal{B}(B^0 \to K^{*+} \pi^-)}$	1.2%	1.8%

7.3 PID requirements

The method used in Section 6.4 to evaluate the PID efficiency has a source of systematic uncertainty, in addition to the statistical uncertainty from the number of simulated events. The method assumes that the calibration PID efficiencies are constant across the range of each efficiency bin in the efficiency histogram. This is an approximation and the associated uncertainty was estimated using two separate methods.

Firstly, simulated calibration track samples exist which are used to perform the same PID efficiency calculation. This gives a value for the PID efficiency on simulated signal events. The PID requirements are then applied to the simulated signal events to get the PID efficiency directly. The difference between the two values gives an approximate value for the uncertainty. This is found to be $\sim 1\%$ on each mode.

However, the above method is sensitive to the reduced statistics of the simulated calibration tracks compared to the data samples. This leads to a potential overestimate of the uncertainty from that method. As a cross check a data driven method was also used.

By applying the PID requirements to the pion or kaon calibration tracks, the efficiency of each requirement on these events was found. Then the calibration tracks are used to calculate the PID requirement efficiency again, but via the efficiency histogram method used in this analysis. The difference between the two values gives an estimate of the systematic due to the binning of the efficiencies. This is performed for a range of PID requirements, the results are shown in Fig. 7.1. This method has the disadvantage that the calibration tracks come from decays with only one charged track, unlike the $B^0_{d,s} \to K^{*\pm} h^{\mp}$ modes.

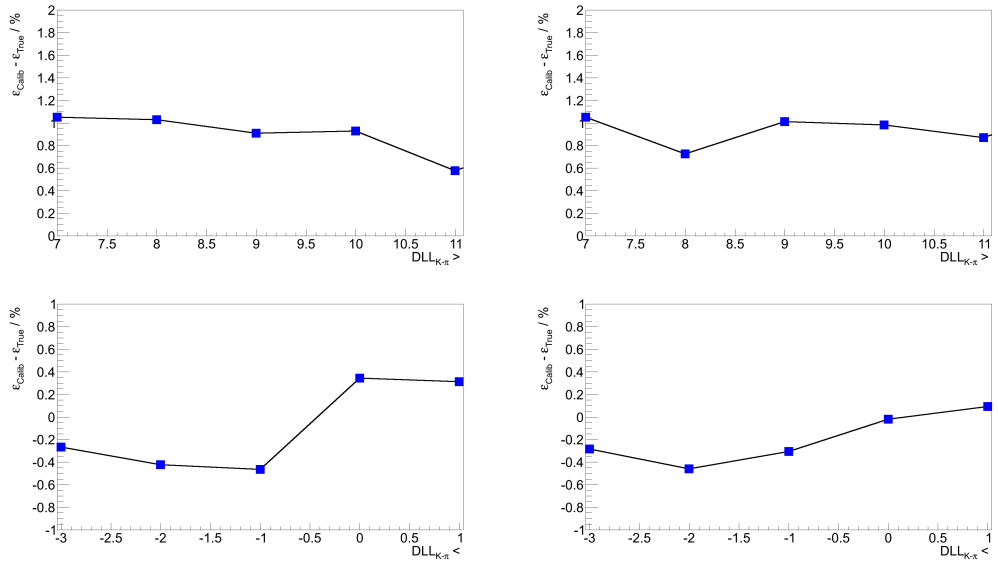

Figure 7.1: PID efficiency differences from data calibration tracks. Top: Kaon DLL cuts. Bottom: Pion DLL cuts: Left: Magnet up. Right: Magnet down.

Comparing the two methods, the systematic uncertainty of $\pm 1\%$ obtained from the first method seems to be a good estimate. This value was also used in the $B^0_{d,s} \to K^0_S h^{\pm} h'^{\mp}$ analysis [1].

7.4 Fit model

The fit model used to extract the yields has several distinct sources of uncertainty. These can be broadly categorised into two types:

1. Coming from known approximations to the physics of the decays.
2. From bias or uncertainty in the construction of the fit model itself.

Linear approximation of non-resonant shapes

As discussed in Chapter 5, the non-resonant decays have their shapes in the $m_{K^0_S\pi}$ variable modelled by linear PDFs. Although the shape fits the simulated events well, the linear shape is an approximation which is only valid for a small range of $m_{K^0_S\pi}$.

In order to estimate the effect of this approximation on the data, the lower bound of the $m_{K^0_S\pi}$ range is varied in the range $550 \to 750$. The resulting selection

and PID efficiencies, as well as the fitted yields, are re-calculated for each value of the lower bound. The number of events prior to the full selection requirements being applied is known. This value can be estimated by dividing the fitted yield by the efficiencies. The percentage variation of this value is used as the percentage systematic uncertainty on the yield, given in Table 7.2.

Table 7.2: Variation in the measured yields due to the linear approximation of the non-resonant modes.

Yield	K^0_S Down-Down	K^0_S Long-Long
$N_{B^0_s \to K^{*\pm} K^{\mp}}$	0.7	0.2
$N_{B^0 \to K^{*\pm} K^{\mp}}$	0.1	0.1
$N_{B^0_s \to K^{*-} \pi^+}$	1.2	0.8
$N_{B^0 \to K^{*+} \pi^-}$	1.8	1.5

S-wave interference

The fit model assumes that there is no component due to interference between the (S-wave) non-resonant term and the K^* resonance. This approximation is only true in the case that the efficiency is symmetric in the cosine of the decay angle, since in that case the interference effect cancels. Non-perfect cancellation of the interference will result in a contribution that peaks in the B candidate mass and has a non-trivial distribution in $m_{K^0_S \pi}$. The efficiency cannot be expected to be symmetric in the cosine of the $K^0_S \pi$ decay angle, since one side corresponds to a slow K^0_S (in the B rest frame) and the other to a slow pion, and therefore this effect could contribute significant uncertainty.

To estimate this systematic uncertainty the decay angle distribution of the K^* is exploited in a method similar to that used the recent analysis on the search for $B^+ \to \phi \pi^+$ decays [72,73]. The decay angle, θ_{K^*}, is defined as the angle between the B and K^0_S candidate momenta in the rest frame of the K^*. The $\cos\theta_{K^*}$ distributions for the $B^0_s \to K^{*\pm} K^{\mp}$ simulated events are shown in Fig. 7.2. The physical distribution should be parabolic, which is shown to be true for values between zero and one. However, the distribution has a significantly different shape for the negative values, due to the lower efficiency on the low p_T bachelor tracks.

As the negative values have a more difficult to parameterise shape this range is not used for the systematic uncertainty measurement. To check that the efficiency effect is negligible in the positive region, the distribution is fitted to a second order polynomial, as shown in Fig. 7.3. The values of the zero and first order coefficients

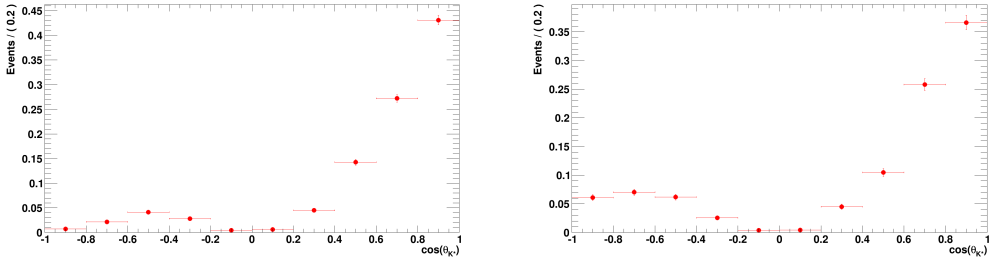

Figure 7.2: Distribution of $\cos\theta_{\mathrm{K}^*}$ for simulated $\mathrm{B}^0_s \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$ events, scaled to unit area. Left: Down-Down. Right: Long-Long.

(p_0 and p_1) are consistent with zero for the $\mathrm{B}^0_s \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$ simulated events, as expected. The fit values for the $\mathrm{B}^0 \to \mathrm{K}^{*+}\pi^-$ simulated events are less consistent with $(p_0, p_1) = (0, 0)$, though still in reasonable agreement, suggesting that the lack of an efficiency correction across this distribution may only be an approximation. However, this sample has lower statistics, and is obtained from re-weighted $\mathrm{B}^0 \to \mathrm{K}^0_{\mathrm{S}}\pi^+\pi^-$ events. In addition, there is no reason to expect significantly different behaviour between K*K and K*π in this variable. Therefore, since the $\mathrm{B}^0_s \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$ distributions show no significant deviation due to efficiency effects, the efficiency correction can be neglected within this fit range.

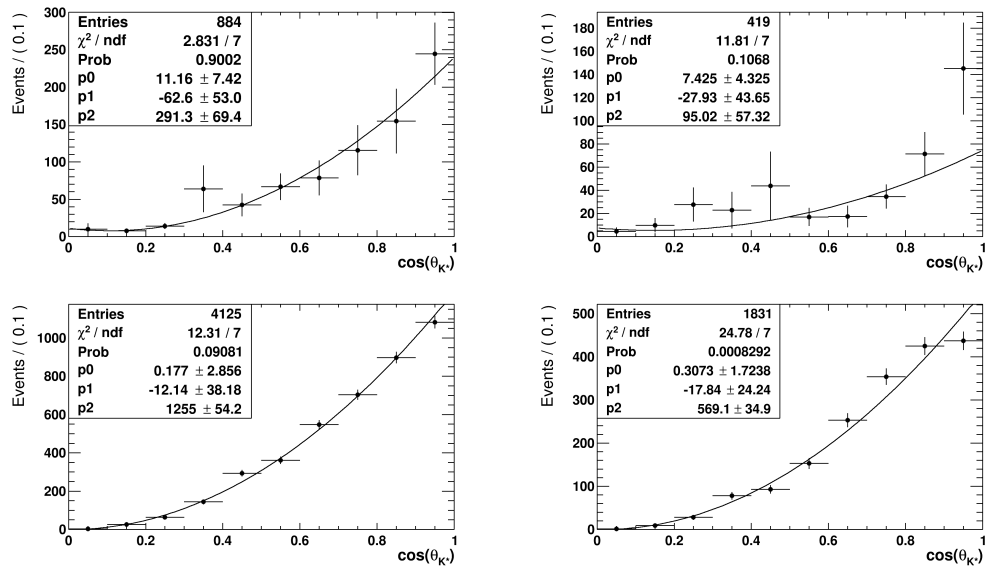

Figure 7.3: Distribution of $\cos\theta_{\mathrm{K}^*}$ for simulated $\mathrm{B}^0_{d,s} \to \mathrm{K}^{*\pm}\mathrm{h}^{\mp}$ events. Left: Down-Down. Right: Long-Long. Top: $\mathrm{B}^0 \to \mathrm{K}^{*+}\pi^-$. Bottom: $\mathrm{B}^0_s \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$.

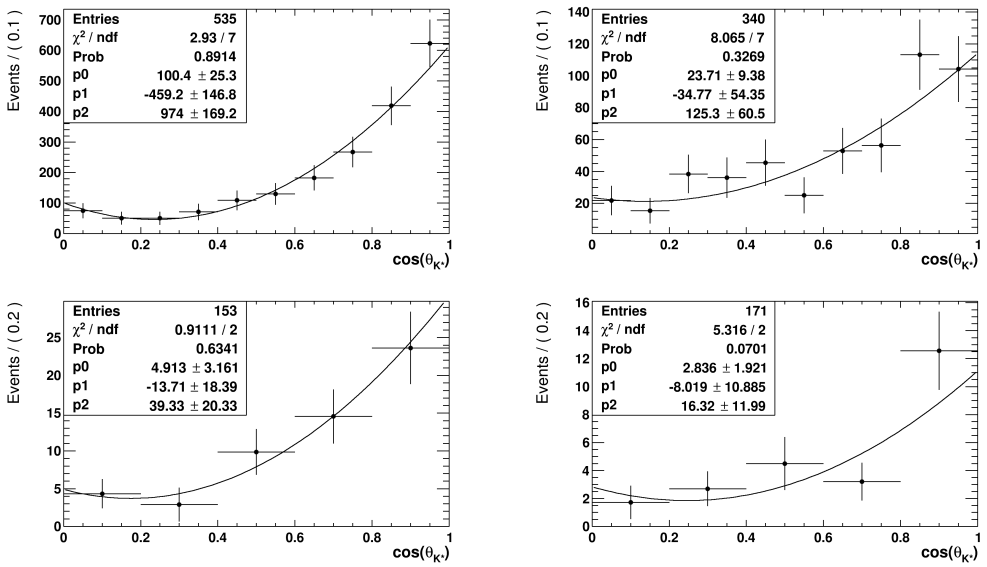

Figure 7.4: Distribution of $\cos\theta_{\mathrm{K}^*}$ for collision $\mathrm{B}^0_{d,s} \to \mathrm{K}^{*\pm}\mathrm{h}^{\mp}$ events. Left: Down-Down. Right: Long-Long. Top: $\mathrm{B}^0 \to \mathrm{K}^{*+}\pi^-$. Bottom: $\mathrm{B}^0_s \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$.

To get the $\cos\theta_{\mathrm{K}^*}$ distribution from the data, the ${}_s\mathcal{P}lot$ weighting method was used. The nominal fit provided the ${}_s\mathcal{W}eights$ to obtain the $\cos\theta_{\mathrm{K}^*}$ distributions for the $\mathrm{B}^0_s \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$ and $\mathrm{B}^0 \to \mathrm{K}^{*+}\pi^-$ models. These distributions are plotted in Fig. 7.4 and the second order polynomial fit results displayed. Recall that the nominal fit separates signal (containing the K^* resonance) from non-resonant $\mathrm{K}^0_{\mathrm{S}}\mathrm{K}^{\pm}\pi^{\mp}$ decays, and therefore we do not expect the (flat) non-resonant contribution to the decay angle distribution to contribute to the obtained distribution. The interference contribution would appear as a linear term in the distribution. To evaluate the systematic uncertainty, first the integrals of the second order polynomial fit to the data distributions were calculated, when all parameters were allowed to float. Then the p_0 and p_1 parameters were fixed to zero and the fit was re-performed with only the p_2 parameter floating. The percentage difference in the integrals between these two fits is quoted as the systematic on the relevant yield.

The measured yields of the $\mathrm{B}^0_s \to \mathrm{K}^{*-}\pi^+$ and $\mathrm{B}^0 \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$ are too small to give a reasonable fit to the distribution. Therefore the same relative uncertainties were applied to these decays as in the corresponding B^0 or B^0_s decay. The resulting relative uncertainties on the yields are given in Table 7.3.

Table 7.3: Measured systematic uncertainty percentages on measured yields from S-wave interference.

Yield	K^0_S Down-Down	K^0_S Long-Long
$N_{B^0_s \to K^{*\pm}K^{\mp}}$	7.8%	13.2%
$N_{B^0 \to K^{*\pm}K^{\mp}}$	7.8%	13.2%
$N_{B^0_s \to K^{*-}\pi^+}$	7.0%	12.0%
$N_{B^0 \to K^{*+}\pi^-}$	7.0%	12.0%

Shape parameters

Many of the PDFs in the fit models have fixed shape parameters. The choice of fixed value comes from fits to simulated events or sideband data, which is subject to an uncertainty from the statistics of the sample used to determine the value.

In order to estimate these uncertainties, the fixed parameter values for the PDFs are varied according to their covariance matrices, which are taken from the fits that determined the values of the parameters. Then the new fixed values were used to re-fit the data and produce new values for the yields. The variations from these fitted yields were taken as the uncertainties, and are given in Table 7.4.

The combinatorial background parameters were not taken from simulated events, but instead from sideband data. The amount of data in the sidebands was extremely small. Therefore the Down-Down K^0_S and Long-Long K^0_S combinatorial PDF parameters were instead floated in separate fits to the data. The resulting covariance matrices were used to perform the variation of the fixed combinatorial parameters.

Fit bias

A subtle effect of the fit models is the potential bias. The fit models may systematically fit higher or lower yields than were actually in the data sample. This may occur in high-statistics fits when there are neglected correlations between variables, or in low-statistics fits due to some intrinsic statistical features. To measure this, simulation studies to toy data samples were performed, using the nominal fitted values as the input parameters to the generation of the toy events.

The residuals of the yields are then plotted which, for an unbiased fit, should be a Gaussian distribution with the width equal to the statistical uncertainty. As examples, the Down-Down K^0_S $B^0_s \to K^{*\pm}K^{\mp}$ and Long-Long Down-Down $B^0 \to K^{*+}\pi^-$ residual distributions are shown in Fig. 7.5. The fitted means of these

Table 7.4: Systematic uncertainties on the signal yields due to the fixing of shape parameters. Uncertainties are divided into the contributions from the signal PDF parameters, combinatorial background parameters, and parameters of backgrounds that arise from specific decay modes. Given here in terms of the absolute number of events, which are added in quadrature for the full systematic uncertainty.

Yield	Signal	Background	Combinatorial
$N^{\mathrm{DD}}_{\mathrm{B}^0_s\to\mathrm{K}^{*\pm}\mathrm{K}^{\mp}}$	0.26	0.37	0.46
$N^{\mathrm{LL}}_{\mathrm{B}^0_s\to\mathrm{K}^{*\pm}\mathrm{K}^{\mp}}$	0.15	0.19	0.41
$N^{\mathrm{DD}}_{\mathrm{B}^0\to\mathrm{K}^{*\pm}\mathrm{K}^{\mp}}$	0.14	0.13	0.43
$N^{\mathrm{LL}}_{\mathrm{B}^0\to\mathrm{K}^{*\pm}\mathrm{K}^{\mp}}$	0.07	0.22	0.43
$N^{\mathrm{DD}}_{\mathrm{B}^0_s\to\mathrm{K}^{*-}\pi^{+}}$	0.09	0.48	0.67
$N^{\mathrm{LL}}_{\mathrm{B}^0_s\to\mathrm{K}^{*-}\pi^{+}}$	0.02	0.35	0.54
$N^{\mathrm{DD}}_{\mathrm{B}^0\to\mathrm{K}^{*+}\pi^{-}}$	0.24	0.55	3.66
$N^{\mathrm{LL}}_{\mathrm{B}^0\to\mathrm{K}^{*+}\pi^{-}}$	0.12	0.58	1.42

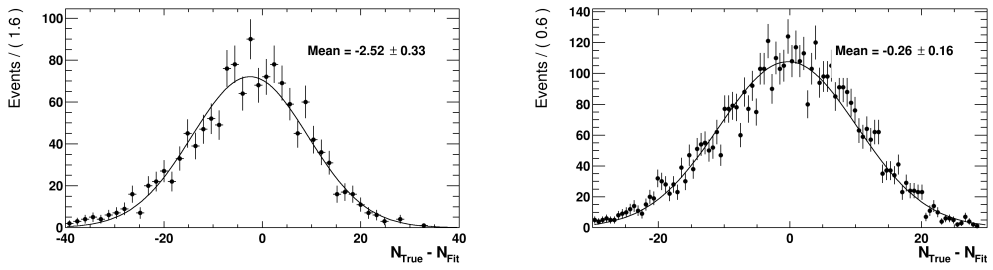

Figure 7.5: Residual distributions from fits to toy events. Left: $\mathrm{B}^0_s \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$ Down-Down. Right: $\mathrm{B}^0 \to \mathrm{K}^{*+}\pi^{-}$ Long-Long.

distributions are detailed in Table 7.5. The final systematic uncertainty on each yield is taken to be the absolute value of the mean plus its uncertainty.

Table 7.5: Fitted means of residuals from fits to toy data samples.

Yield	K^0_S Down-Down	K^0_S Long-Long
$N_{B^0_s \to K^{*\pm} K^{\mp}}$	-2.5 ± 0.3	-2.6 ± 0.2
$N_{B^0 \to K^{*\pm} K^{\mp}}$	-0.7 ± 0.1	-0.6 ± 0.1
$N_{B^0_s \to K^{*-} \pi^+}$	0.0 ± 0.1	0.0 ± 0.1
$N_{B^0 \to K^{*+} \pi^-}$	-0.6 ± 0.3	-0.3 ± 0.2

7.5 Summary

Here the systematic uncertainties are summarised and the totals given. The systematics contributing to the observation significance are given in Tables 7.6-7.7. These are given in terms of the absolute yield of the relevant decay. The systematics contributing to the branching fraction measurements and limits, are given in Tables 7.8-7.9 in terms of the relative branching fractions. The uncertainties on the efficiency measurements are quoted as additional sources of systematic uncertainty on the branching fractions.

Table 7.6: Summary of systematic uncertainties on the Long-Long K^0_S yields. Given in terms of absolute number of events measured.

Contribution	$N_{B^0_s \to K^{*\pm} K^{\mp}}$	$N_{B^0 \to K^{*\pm} K^{\mp}}$	$N_{B^0_s \to K^{*-} \pi^+}$	$N_{B^0 \to K^{*+} \pi^-}$
Shape parameters	0.5	0.5	0.6	1.5
Fit bias	2.8	0.7	0.1	0.5
S-wave interference	5.3	0.1	0.6	9.6
Linear approximation	0.2	0.1	0.8	1.5
Total	6.0	0.9	1.2	9.8

Table 7.7: Summary of systematic uncertainties on the Down-Down K^0_S yields. Given in terms of absolute number of events measured.

Contribution	$N_{B^0_s \to K^{*\pm} K^{\mp}}$	$N_{B^0 \to K^{*\pm} K^{\mp}}$	$N_{B^0_s \to K^{*-} \pi^+}$	$N_{B^0 \to K^{*+} \pi^-}$
Shape parameters	0.6	0.5	0.8	3.7
Fit bias	2.8	0.8	0.1	0.9
S-wave interference	4.9	0.3	1.6	11.5
Linear approximation	0.7	0.1	1.2	1.8
Total	5.7	1.0	2.2	12.2

Table 7.8: Summary of systematic uncertainties on the Long-Long K^0_S branching fractions, relative to the $B^0 \to K^{*+}\pi^-$ branching fraction. The fit systematic is from the combined systematic contributions detailed in Table 7.6.

Contribution	$\frac{\mathcal{B}(B^0_s \to K^{*\pm} K^{\mp})}{\mathcal{B}(B^0 \to K^{*+}\pi^-)}$	$\frac{\mathcal{B}(B^0 \to K^{*\pm} K^{\mp})}{\mathcal{B}(B^0 \to K^{*+}\pi^-)}$	$\frac{\mathcal{B}(B^0_s \to K^{*-}\pi^+)}{\mathcal{B}(B^0 \to K^{*+}\pi^-)}$
Acceptance	0.01	0.4×10^{-4}	0.003
Trigger	0.03	0.5×10^{-4}	0.003
Selection	0.08	2.1×10^{-4}	0.012
PID	0.04	1.1×10^{-4}	0.006
Fit	0.27	110×10^{-4}	0.055
f_s/f_d	0.10	–	0.013
Total	0.30	0.011	0.058

Table 7.9: Summary of systematic uncertainties on the Down-Down K^0_S branching fractions, relative to the $B^0 \to K^{*+}\pi^-$ branching fraction. The fit systematic is from the combined systematic contributions detailed in Table 7.7.

Contribution	$\frac{\mathcal{B}(B^0_s \to K^{*\pm} K^{\mp})}{\mathcal{B}(B^0 \to K^{*+}\pi^-)}$	$\frac{\mathcal{B}(B^0 \to K^{*\pm} K^{\mp})}{\mathcal{B}(B^0 \to K^{*+}\pi^-)}$	$\frac{\mathcal{B}(B^0_s \to K^{*-}\pi^+)}{\mathcal{B}(B^0 \to K^{*+}\pi^-)}$
Acceptance	0.010	0.2×10^{-3}	0.008
Trigger	0.019	0.8×10^{-3}	0.010
Selection	0.045	0.8×10^{-3}	0.020
PID	0.028	0.5×10^{-3}	0.011
Fit	0.126	5.7×10^{-3}	0.051
f_s/f_d	0.080	–	0.031
Total	0.16	5.8×10^{-3}	0.065

8 Results

8.1 Introduction

This chapter describes the final combinations of the previous parts of the analysis into a measurement of the relative branching fractions. Firstly, in Section 8.2 the significances of the measured yields of the various decays are calculated. Then, based on the observed significance, the relative branching fraction is reported as either an upper limit or as a central value with statistical and systematic uncertainties. The methods for both of these are detailed, and the results are presented in Sections 8.3-8.4.

8.2 Observation significances

The yields of the three signal modes determined from fits to the data samples are given in Tables 5.8-5.9. The significance of these yields when compared to the null hypothesis (zero events) remains to be determined. As the fitting technique used the maximum likelihood method, a useful estimation of the significance comes from the likelihood ratio test which is used extensively in particle physics [74].

Likelihood ratio test

During the nominal fit to the data, the maximum likelihood is found via minimisation of the negative log-likelihood, $-L = -\ln(\mathcal{L})$, where $\mathcal{L}$ is the likelihood. When the value of one of the yields is fixed before minimisation, the fit has one less degree of freedom and therefore returns a value of L which is smaller than the one returned by the nominal fit. A useful relation between the likelihood and the often used χ^2 distribution is

$$\chi^2 \approx -2\ln\left(\frac{\mathcal{L}_0}{\mathcal{L}}\right), \tag{8.1}$$

where $\mathcal{L}_0$ is the likelihood returned by the fit when the parameter of interest is fixed to a certain value [75]. The value of $\ln\left(\frac{\mathcal{L}_0}{\mathcal{L}}\right)$ is often called the *Delta Log-Likelihood*, $\Delta\ln(\mathcal{L})$. As the difference between the fits giving $\mathcal{L}$ and $\mathcal{L}_0$ is only a single degree of freedom, the significance (in standard deviations σ) of the measurement compared to the fixed value is given by $\sigma = \sqrt{\chi^2}$. In this analysis the comparison is with the case where the yield of a species is fixed to be zero.

Significances

The values of $-\Delta\ln(\mathcal{L})$ when each K^0_S yield for each signal mode was fixed to various values are shown in Figs. 8.1-8.3. In these cases, the dashed red line shows the raw $-\Delta\ln(\mathcal{L})$ values. The solid blue curve shows $-\Delta\ln(\mathcal{L})$ after being convolved with a Gaussian of width equal to the total fit systematic uncertainties on the measured yields, given in Tables 7.6-7.7. This allows the systematic uncertainty on the measured yields to be taken into account when calculating the final observation significances. The statistical and total significances for each signal mode and K^0_S sample are given separately in Tables 8.1-8.2.

Table 8.1: Long-Long K^0_S sample observation significances of signal modes. The total significance includes the systematic uncertainty.

Decay mode	Statistical significance (σ)	Total significance (σ)
$\mathrm{B}^0_s \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$	7.8	4.7
$\mathrm{B}^0 \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$	0.0	0.0
$\mathrm{B}^0_s \to \mathrm{K}^{*-}\pi^+$	1.3	1.2

Until now, the K^0_S samples have been treated as separate measurements. However, both samples come from the same underlying decay process and it makes

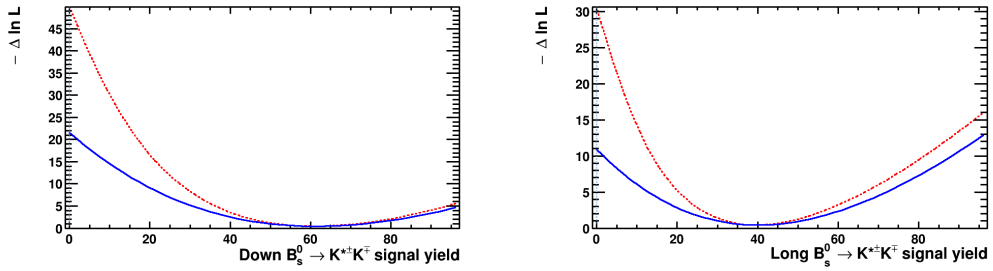

Figure 8.1: Negative delta log-likelihood curves for the $B_s^0 \to K^{*\pm}K^{\mp}$ mode. Red curve is the raw curve from the fit model. Blue is after the relevant systematic uncertainty has been accounted for. Left: Down-Down. Right: Long-Long.

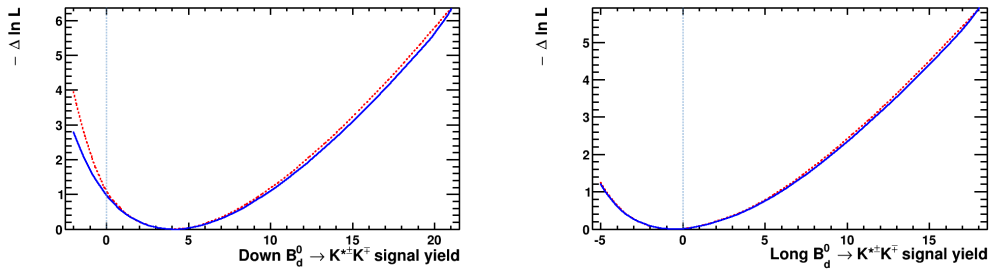

Figure 8.2: Negative delta log-likelihood curves for the $B^0 \to K^{*\pm}K^{\mp}$ mode. Red curve is the raw curve from the fit model. Blue is after the relevant systematic uncertainty has been accounted for. Left: Down-Down. Right: Long-Long.

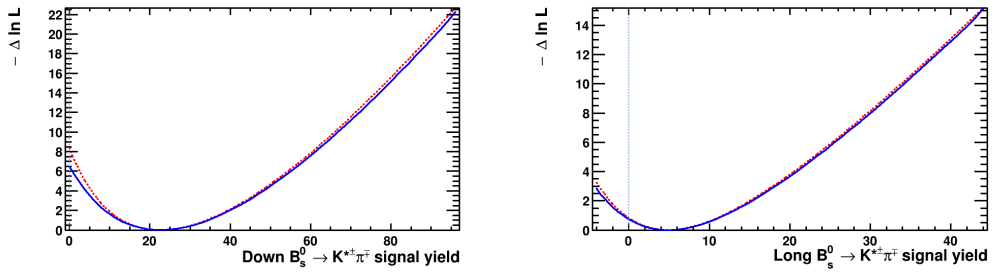

Figure 8.3: Negative delta log-likelihood curves for the $B_s^0 \to K^{*-}\pi^{+}$ mode. Red curve is the raw curve from the fit model. Blue is after the relevant systematic uncertainty has been accounted for. Left: Down-Down. Right: Long-Long.

Table 8.2: Down-Down K^0_S sample observation significances of signal modes. The total significance includes the systematic uncertainty.

Decay mode	Statistical significance (σ)	Total significance (σ)
$B^0_s \to K^{*\pm}K^{\mp}$	10.0	6.6
$B^0 \to K^{*\pm}K^{\mp}$	1.5	1.4
$B^0_s \to K^{*-}\pi^+$	4.1	3.6

Table 8.3: Combined K^0_S sample observation significances of signal modes. The total significance includes the systematic uncertainty. Combination not performed for the $B^0 \to K^{*\pm}K^{\mp}$.

Decay mode	Statistical significance (σ)	Total significance (σ)
$B^0_s \to K^{*\pm}K^{\mp}$	12.5	7.8
$B^0_s \to K^{*-}\pi^+$	3.9	3.4

sense to combine the statistical power of the two measurements into a single observation significance for a decay. The likelihood curves of the two yields cannot be simply added, as the two axes have different meanings. Instead, the value of the χ^2 for each K^0_S sample at the null hypothesis was found and added together to give

$$\chi^2_{K^0_S} = \chi^2_{DD} + \chi^2_{LL}. \tag{8.2}$$

The value of $\chi^2_{K^0_S}$ was used to give a p-value, under the assumption that there are two degrees of freedom associated with the χ^2. Then the equivalent χ^2 value was found that gives the same p-value for only one degree of freedom. This χ^2 value was then used to calculate the total observation significance of the two K^0_S samples, which are listed in Table 8.3.

For these results there are several qualities to note. Firstly, the nominal fit returned a negative value for the yield of Long-Long $B^0 \to K^{*\pm}K^{\mp}$ events. The observation significance is therefore zero and the combined significance is not performed. This does not affect the final conclusions of this analysis as both of the Long-Long and Down-Down yields for this mode were consistent with zero anyway. A total significance for these modes is therefore also not calculated, but a relative branching fraction limit is determined in Section 8.3.

Secondly, the $B^0_s \to K^{*-}\pi^+$ combined significance is slightly lower than the Down-Down K^0_S only significance. This is due to the low yield of the Long-Long K^0_S

measurement and is to be expected when a very low statistics sample is combined with a larger one. Again, as the combined value is still $\geq 3\,\sigma$ making no change to the final conclusion, no further analysis is undertaken. For the $B_s^0 \to K^{*\pm}K^{\mp}$ mode a final combined significance of $7.8\,\sigma$ is found. As the $B_s^0 \to K^{*\pm}K^{\mp}$ and $B_s^0 \to K^{*-}\pi^+$ decays have significances $\geq 3\,\sigma$, central values for the relative branching are calculated in Section 8.4.

8.3 Relative branching fraction limits

As there is no evidence for the $B^0 \to K^{*\pm}K^{\mp}$ decay, a limit was calculated on the relative branching fraction. The measured efficiencies, nominal fit results and master formula for the relative branching fraction, Eq. 8.3, were used to calculate the yields of $B^0 \to K^{*\pm}K^{\mp}$ required to give specific branching fractions. The yields in the fit were then fixed to these values, for the different K_S^0 samples separately, so that delta log-likelihood curves could be created with the relative branching fraction as the x-axis.

As before, the systematic uncertainties needed to be taken into account. However, as the relative branching fraction was being calculated for the likelihood curve, the full set of systematics were applied from Tables 7.8-7.9, rather than just the yield systematics. The systematics which are uncorrelated between the K_S^0 samples were applied by convolving the separate likelihood curves with Gaussian functions of the relevant widths. The resulting delta log-likelihood curves were then added and the correlated systematic uncertainties applied in a further convolution. For this analysis, the PID and acceptance systematics were considered to be correlated. The final likelihood curve, plotted as a function of the relative branching fraction is shown in Fig. 8.4.

In Bayesian statistics, the posterior probability after having accounted for the experimental evidence is proportional to the likelihood multiplied by the prior probability. In this analysis, a uniform prior probability is assumed, so that no branching fraction is preferred over any other. This means that by integrating the area under the likelihood curve for the physical branching fraction values (> 0), upper confidence limits can be calculated. The values of the relative branching fraction which gave 90% and 95% upper confidence limits were found and are shown in Fig. 8.4.

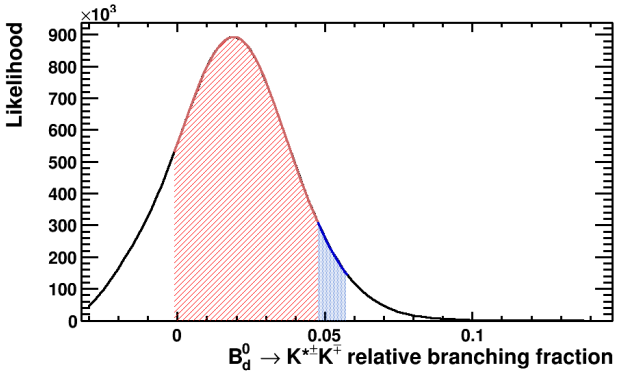

Figure 8.4: Combined Long-Long and Down-Down likelihood curve for the $B^0 \to K^{*\pm}K^{\mp}$ mode as a function of the relative branching fraction. Red and blue areas show the 90% (< 0.05) and 95% (< 0.06) confidence limits respectively. Systematic uncertainties are convolved into this curve.

8.4 Relative branching fractions

The $B^0_s \to K^{*-}\pi^+$ and $B^0_s \to K^{*\pm}K^{\mp}$ relative branching fractions are calculated here. The master formula for the relative branching fractions is

$$\frac{\mathcal{B}\left(B^0_s \to K^{*\pm}K^{\mp}\right)}{\mathcal{B}\left(B^0 \to K^{*+}\pi^-\right)} = \frac{f_d}{f_s}.R^{\mathrm{Acc}}.R^{\mathrm{Sel}}.R^{\mathrm{PID}}.\frac{N_{B^0_s \to K^{*\pm}K^{\mp}}}{N_{B^0 \to K^{*+}\pi^-}}, \tag{8.3}$$

where R^{α} is the ratio of the $B^0 \to K^{*+}\pi^-$ efficiency to the $B^0_s \to K^{*\pm}K^{\mp}$ efficiency. However, the efficiency ratios calculated in Chapter 6 may differ based on the magnet polarity and on the K^0_S type (Down-Down or Long-Long). The integrated luminosity for each of the magnet polarities is known and was used to perform the weighted average

$$R_{Avg} = \frac{\mathcal{L}_{\mathrm{Up}}R_{\mathrm{Up}} + \mathcal{L}_{\mathrm{Down}}R_{\mathrm{Down}}}{\mathcal{L}_{\mathrm{Up}} + \mathcal{L}_{\mathrm{Down}}}, \tag{8.4}$$

to obtain the overall ratio. This average was performed once, on the combined ratio

$$R^{\mathrm{Total}} = R^{\mathrm{Acc}}R^{\mathrm{Sel}}R^{\mathrm{PID}}. \tag{8.5}$$

A summary of the values entering the branching fraction calculations are given in Table 8.4.

Additionally, the different K^0_S sample relative branching fractions also needed to be combined. The branching ratios, prior to the multiplication by the f_s/f_d hadronisation fraction, for the K^0_S Long-Long and K^0_S Down-Down samples were

Table 8.4: Summary of measured efficiency ratios, yields and physical constants used in the calculation of the branching fractions. Ratios are averaged for the different magnet polarities and uncertainties are only given for the statistical source of uncertainty from the measured yields.

Value	Down-Down K^0_S	Long-Long K^0_S
$R^{\mathrm{Acc}}_{\mathrm{B}^0_s\to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}}$	1.00	1.00
$R^{\mathrm{Acc}}_{\mathrm{B}^0\to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}}$	1.00	1.00
$R^{\mathrm{Acc}}_{\mathrm{B}^0_s\to \mathrm{K}^{*-}\pi^{+}}$	1.01	1.01
$R^{\mathrm{Sel}}_{\mathrm{B}^0_s\to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}}$	0.87	0.89
$R^{\mathrm{Sel}}_{\mathrm{B}^0\to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}}$	0.88	0.92
$R^{\mathrm{Sel}}_{\mathrm{B}^0_s\to \mathrm{K}^{*-}\pi^{+}}$	0.97	0.93
$R^{\mathrm{PID}}_{\mathrm{B}^0_s\to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}}$	1.08	1.04
$R^{\mathrm{PID}}_{\mathrm{B}^0\to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}}$	1.07	1.03
$R^{\mathrm{PID}}_{\mathrm{B}^0_s\to \mathrm{K}^{*-}\pi^{+}}$	1.00	1.01
$N_{\mathrm{B}^0_s\to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}}$	62 ± 10	40 ± 8
$N_{\mathrm{B}^0\to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}}$	4 ± 3	0 ± 4
$N_{\mathrm{B}^0_s\to \mathrm{K}^{*-}\pi^{+}}$	23 ± 8	5 ± 4
$N_{\mathrm{B}^0\to \mathrm{K}^{*+}\pi^{-}}$	165 ± 16	80 ± 10
f_s/f_d	0.259	0.259 [76]

found separately, and then the weighted average

$$\mathcal{B}\left(\text{Avg}\right) = \frac{w_{\text{LL}}\mathcal{B}\left(\text{LL}\right) + w_{\text{DD}}\mathcal{B}\left(\text{DD}\right)}{w_{\text{LL}} + w_{\text{DD}}}, \tag{8.6}$$

was made. Here w_{LL} and w_{DD} are taken from the statistical uncertainty on the final branching ratios, $\sigma_{\mathcal{B}}$, as

$$w = \frac{1}{\sigma_{\mathcal{B}}^2}. \tag{8.7}$$

The systematic uncertainties on these calculations are summarised in Tables 7.8-7.9. For the calculation of the $\text{B}_s^0 \to \text{K}^{*-}\pi^+$ and $\text{B}_s^0 \to \text{K}^{*\pm}\text{K}^{\mp}$ the f_s/f_d uncertainty is an additional source of systematic uncertainty, which is correlated between the Down-Down and Long-Long relative branching fractions. As in the estimation of the relative branching fraction limits, the PID and acceptance systematics are assumed to be fully correlated. The final combined systematic uncertainties were performed using an uncertainty propagation package [77]. A summary of the final relative branching fractions and the naively averaged theoretical predictions from Table 1.4 is given in Table 9.1. Using the HFAG average of $\mathcal{B}\left(\text{B}^0 \to \text{K}^{*+}\pi^-\right) = (8.5 \pm 0.7) \times 10^{-6}$ the absolute branching fractions and limits were calculated and compared against the theoretical predictions in Table 9.2. The uncertainty on the measured value of $\mathcal{B}\left(\text{B}^0 \to \text{K}^{*+}\pi^-\right)$ is used as an additional systematic uncertainty.

9 Conclusion

9.1 Summary of Results

The $B^0_{d,s} \to K^{*\pm}h^{\mp}$ decays were investigated in $1.0\,\text{fb}^{-1}$ of data at a centre-of-mass energy of 7 TeV, collected by LHCb during 2011. The decay $B^0_s \to K^{*\pm}K^{\mp}$ is observed for the first time, with a combined statistical and systematic significance of 7.8σ. Evidence for the $B^0_s \to K^{*-}\pi^+$ decay is presented, with a significance of 3.4σ. No significant evidence for the $B^0 \to K^{*\pm}K^{\mp}$ decay is found. A branching fraction, or limit, relative to the previously measured $B^0 \to K^{*+}\pi^-$ branching fraction is given for each of these three decays, see Table 9.1. Multiplication by the previously measured $B^0 \to K^{*+}\pi^-$ branching fraction gives the values and limit in terms of the absolute branching fractions, see Table 9.2.

Table 9.1: Measurements, limits and naively averaged theoretical predictions of the relative branching fractions. For the measured values, the first uncertainty is statistical and the second systematic. The limit on the $B^0 \to K^{*\pm}K^{\mp}$ measured value is the 95% confidence limit.

Relative branching fraction	Average theoretical prediction	Measured value
$\frac{\mathcal{B}\left(B_s^0 \to K^{*\pm}K^{\mp}\right)}{\mathcal{B}(B^0 \to K^{*+}\pi^-)}$	1.7 ± 0.3	$1.49 \pm 0.22 \pm 0.18$
$\frac{\mathcal{B}\left(B^0 \to K^{*\pm}K^{\mp}\right)}{\mathcal{B}(B^0 \to K^{*+}\pi^-)}$	0.018 ± 0.005	< 0.06
$\frac{\mathcal{B}\left(B_s^0 \to K^{*-}\pi^+\right)}{\mathcal{B}(B^0 \to K^{*+}\pi^-)}$	0.90 ± 0.09	$0.39 \pm 0.13 \pm 0.05$

Table 9.2: Measurements, limits and theoretical predictions of the absolute branching fractions in units of ($\times 10^{-6}$). For the measured values, the first uncertainty is statistical and the second systematic. The limit on the $B^0 \to K^{*\pm}K^{\mp}$ measured value is the 95% confidence limit. The three different theoretical predictions are discussed in Section 1.5.

Branching fraction	QCDF [35]	pQCD [41]	Effective Hamiltonian [42]	Measured value
$\mathcal{B}\left(B_s^0 \to K^{*\pm}K^{\mp}\right)$	24 ± 10	12 ± 3	16 ± 3	$12.7 \pm 1.9 \pm 1.9$
$\mathcal{B}\left(B^0 \to K^{*\pm}K^{\mp}\right)$	0.15 ± 0.04	-	-	< 0.5
$\mathcal{B}\left(B_s^0 \to K^{*-}\pi^{+}\right)$	$7.8^{+0.4+0.5}_{-0.7-0.7}$	$7.6^{+2.9+0.4+0.5}_{-2.2-0.5-0.3}$	$7.2^{+5.6+0.7}_{-2.2-0.5}$	$3.3 \pm 1.1 \pm 0.5$

9.2 Comparison with predictions

As shown in Table 9.2, three separate theoretical models are used for comparison with these results. Predictions from QCD factorization, Perturbative QCD and a six quark effective Hamiltonian model are considered. The measured value of $\mathcal{B}\left(\mathrm{B}_s^0 \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}\right)$ is compatible with all three models. The branching fraction of the $\mathrm{B}^0 \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$ decay was only calculated in the QCD Factorization case and is compatible with the limit set by this analysis. The value measured for $\mathcal{B}\left(\mathrm{B}_s^0 \to \mathrm{K}^{*-}\pi^{+}\right)$ is lower than expected from the theoretical predictions. However, given that the evidence for the $\mathrm{B}_s^0 \to \mathrm{K}^{*-}\pi^{+}$ observation is low due to the dominant statistical uncertainty, no discrepancy with the theoretical predictions is here claimed.

9.3 Future analyses

Looking to future analyses of these modes, probably the most important question is, how can the uncertainties be reduced? This analysis uses the full 2011 LHCb data sample, which is approximately $1.0\,\mathrm{fb}^{-1}$. However, in 2012 approximately $2.1\,\mathrm{fb}^{-1}$ were recorded at LHCb at a centre-of-mass energy of 8 TeV, leading to a data sample about three times larger when both are combined. The $\mathrm{b\overline{b}}$ cross-section is increased by $\sim 8/7$ due the larger centre-of-mass energy. Additionally, the trigger efficiency for modes containing a K_S^0 was improved. If an analysis could use both samples together, or even just 2012 on its own, there should be a corresponding decrease in the statistical uncertainties ($\sim \times 2$). This is the dominant source of uncertainty for the two remaining unobserved decays in this analysis.

If the $\mathrm{B}_s^0 \to \mathrm{K}^{*-}\pi^{+}$ decay has a true branching fraction at the value measured in this analysis, then using the combined 2011 and 2012 data samples could potentially yield an observation. Additionally, the branching fraction limit on the $\mathrm{B}^0 \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$ decay could be significantly lowered towards its theoretical value, if the predictions are true. *CP*-violation studies with the $\mathrm{B}_{d,s}^0 \to \mathrm{K}^{*\pm}\mathrm{h}^{\mp}$ modes are also very desirable. In the case of the $\mathrm{B}_{d,s}^0 \to \mathrm{K}^{*\pm}\pi^{\mp}$ flavour specific decays, this can be done with a time-integrated study that would not require flavour-tagging. However, as the $\mathrm{B}_s^0 \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$ decay is not flavour specific, a time-dependent and flavour-tagged study would be necessary.

In the case of the observed $\mathrm{B}_s^0 \to \mathrm{K}^{*\pm}\mathrm{K}^{\mp}$, the statistical and systematic uncertainties are very similar in size. A reduction in the statistical uncertainty would still be useful as it might bring the total experimental uncertainty well below the theoretical uncertainty. However, systematic uncertainties are often harder to

reduce as they are not necessarily sensitive to the increasing statistics. For this branching fraction measurement, the dominant sources of systematic uncertainty come from the S-wave interference, fit bias and the f_s/f_d hadronisation fraction, see Tables 7.6-7.9.

The S-wave interference systematic comes from the assumption that the S-wave contribution and K^* resonance do not interfere significantly when making the fit model. The most effective way to reduce or remove this systematic would be to do a full Dalitz plot analysis, which has the added benefit of measuring a larger range of intermediate resonances than this one.

The overall fit bias can arise due to many different effects. However, in this case it seems likely that the bias is at least partially due to the small statistics of the data sample. This can be seen from the fit bias of the higher yield decay $B^0 \to K^{*+}\pi^-$, which was proportionally smaller despite coming from an extremely similar fit model. Therefore there is hope that should a future analysis attempt the same measurement on the larger data samples, the systematic uncertainty from the fit bias should be smaller.

A

Dalitz plots

A.1 Three-body kinematics

The Dalitz plot [78] is an exceptionally useful analysis technique for examining three body decays and their intermediate resonances. It allows the kinematics of three-body decays to be used to distinguish between the resonances.

Firstly, consider the decay of a B meson to three daughter particles named 1, 2 and 3. The four momentum of a daughter particle i is defined as p_i, and its mass is m_i. The combinations $p_{ij} = p_i + p_j$ and $m_{ij}^2 = p_{ij}^2$ can be defined for each pair of daughters. From the conservation of energy and momentum, this leads to

$$m_{\mathrm{B}}^2 + m_1^2 + m_2^2 + m_3^2 = m_{12}^2 + m_{13}^2 + m_{23}^2. \tag{A.1}$$

Then the quantities m_{ij}^2 can be re-written in terms of the B and remaining daughter as

$$m_{ij}^2 = (p_{\mathrm{B}} - p_k)^2 = m_{\mathrm{B}}^2 + m_k^2 - 2m_{\mathrm{B}}E_k, \tag{A.2}$$

where E_k is the energy of the particle k in the rest frame of the B meson [4].

A.2 Dalitz plots

The decay rate, $d\Gamma$, of the three-body decay can be written in terms of the m_{ij}^2 variables as

$$d\Gamma = \frac{1}{(2\pi)^3 32 m_{\mathrm{B}}^3} |\mathcal{M}|^2 \, dm_{ij}^2 dm_{jk}^2. \tag{A.3}$$

Here $\mathcal{M}$ is the scattering matrix element of the decay.

A scatter plot of any two of the m_{ij}^2 variables is called a Dalitz plot. An example Dalitz plot from the LHCb analysis of the $\mathrm{D}^+ \to \mathrm{K}^-\mathrm{K}^+\pi^+$ decay is given in Fig. A.1. The non-uniformity in the plot is due to the intermediate resonances of the decay. This appears mathematically in Eq. A.3 when the matrix element depends on the kinematics of the decay. The type of structure in the plot gives information about the angular momentum of the resonance. For example, a vector (spin-1) resonance like the $\overline{\mathrm{K}}^*(892)^0$ in Fig. A.1 appears as a two lobed band in the relevant axis.

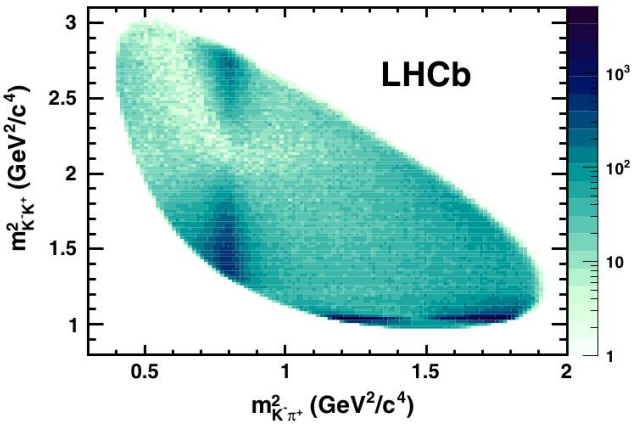

Figure A.1: Dalitz plot of the $\mathrm{D}^+ \to \mathrm{K}^-\mathrm{K}^+\pi^+$ decay using $35\,\mathrm{pb}^{-1}$ of data taken at LHCb during 2010. The vertical $\overline{\mathrm{K}}^*(892)^0$ and horizontal $\phi(1020)$ resonance bands are clearly visible. Reproduced from [79].

Bibliography

[1] LHCb collaboration, *Branching fraction measurements of $B^0_{d,s}$ decays to $K^0_S hh'$ final states, including first observation of $B^0_s \to K_S K\pi$*, LHCb-CONF-2012-023.

[2] CMS Collaboration, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, arXiv:1207.7235.

[3] ATLAS Collaboration, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, arXiv:1207.7214.

[4] J. Beringer et al. (Particle Data Group)Phys. Rev. **D86** (2012) 010001.

[5] ATLAS collaboration, G. Aad *et al.*, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, Tech. Rep. arXiv:1307.1432. CERN-PH-EP-2013-102, CERN, Geneva, Jul, 2013.

[6] CMS Collaboration, S. Chatrchyan *et al.*, *Study of the mass and spin-parity of the higgs boson candidate via its decays to z boson pairs. on the mass and spin-parity of the higgs boson candidate via its decays to z boson pairs*, Phys. Rev. Lett. **110** (2012) 081803. 25 p.

[7] J. C. Maxwell, *A Dynamical Theory of the Electromagnetic Field*, Philosophical Transactions of the Royal Society of London **155** (1865) 459.

[8] A. Einstein, *Relativity: The special and general theory*, H. Holt and Company, 1916.

[9] R. P. Feynman, F. B. Morinigo, and W. G. Wagner, *Feynman lectures on gravitation*, Addison-Wesley, Reading, MA, 1995.

[10] E. Noether, *Invariante Variationsprobleme*, Nachr. Ges. Wiss. Göttingen (1918) 235.

[11] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579.

[12] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[13] A. Salam, *Eighth Nobel Symposium (Stockholm: Almquvist and Wiksells)*, Elementary particle Physics (1968).

[14] K. Anikeev *et al.*, *B physics at the Tevatron: Run II and beyond*, `arXiv:hep-ph/0201071`.

[15] I. J. R. Aitchison and A. J. G. Hey, *Gauge Theories in Particle Physics - Volume 2*, Taylor and Francis Group, 2004.

[16] BABAR, P. F. Harrison and H. R. Quinn, *The BABAR physics book: Physics at an asymmetric* B *factory*, .

[17] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531.

[18] M. Kobayashi and T. Maskawa, *CP violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973) 652.

[19] L.-L. Chau and W.-Y. Keung, *Comments on the parametrization of the Kobayashi-Maskawa matrix*, Phys. Rev. Lett. **53** (1984) 1802.

[20] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[21] J. S. Schwinger, *Spin, statistics, and the TCP theorem*, Proc. Nat. Acad. Sci. **44** (1958) 223.

[22] T. D. Lee and C. N. Yang, *Question of parity conservation in weak interactions*, Phys. Rev. **104** (1956) 254.

[23] C. S. Wu *et al.*, *Experimental test of parity conservation in beta decay*, Phys. Rev. **105** (1957) 1413.

[24] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2 π decay of the K_2^0 meson*, Phys. Rev. Lett. **13** (1964) 138.

[25] BaBar, B. Aubert *et al.*, *Observation of CP violation in the* B^0 *meson system*, Phys. Rev. Lett. **87** (2001) 091801, `arXiv:hep-ex/0107013`.

[26] Belle Collaboration, K. Abe *et al.*, *Observation of large CP violation in the neutral B meson system*, Phys. Rev. Lett. **87** (2001) 091802, arXiv:hep-ex/0107061.

[27] BaBar Collaboration, J. Lees *et al.*, *Observation of Time Reversal Violation in the B^0 Meson System*, Phys. Rev. Lett. **109** (2012) 211801, arXiv:1207.5832.

[28] A. J. Buras, M. E. Lautenbacher, and G. Ostermaier, *Waiting for the top quark mass, $K^+ \to \pi^+$ neutrino anti-neutrino, B(s)0 - anti-B(s)0 mixing and CP asymmetries in B decays*, Phys. Rev. **D50** (1994) 3433, arXiv:hep-ph/9403384.

[29] C. Jarlskog, *Commutator of the quark mass matrices in the Standard Electroweak model and a measure of maximal CP violation*, Phys. Rev. Lett. **55** (1985) 1039.

[30] NA48, V. Fanti *et al.*, *A new measurement of direct CP violation in two pion decays of the neutral kaon*, Phys. Lett. **B465** (1999) 335, arXiv:hep-ex/9909022.

[31] KTeV, A. Alavi-Harati *et al.*, *Observation of direct CP violation in $K_{S,L} \to \pi\pi$ decays*, Phys. Rev. Lett. **83** (1999) 22, arXiv:hep-ex/9905060.

[32] BaBar, B. Aubert *et al.*, *Observation of direct CP violation in $B^0 \to K\pi$ decays*, Phys. Rev. Lett. **93** (2004) 131801, arXiv:hep-ex/0407057.

[33] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation and the B_s^0 meson decay width difference with $B_s^0 \to J/\psi K^+K^-$ and $B_s^0 \to J/\psi\pi^+\pi^-$ decays*, Phys. Rev. **D87** (2013) 112010, arXiv:1304.2600.

[34] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of B-Hadron, C-Hadron, and tau-lepton properties as of early 2012*, arXiv:1207.1158.

[35] H.-Y. Cheng and C.-K. Chua, *QCD Factorization for Charmless Hadronic B_s Decays Revisited*, Phys. Rev. **D80** (2009) 114026, arXiv:0910.5237.

[36] M. Antonelli *et al.*, *Flavor Physics in the Quark Sector*, Phys. Rept. **494** (2010) 197, arXiv:0907.5386.

[37] M. Ciuchini, M. Pierini, and L. Silvestrini, *Hunting the CKM weak phase with time-integrated Dalitz analyses of B(s) $\to$ K K pi and B(s) $\to$ K pi pi decays*, Phys. Lett. **B645** (2007) 201, arXiv:hep-ph/0602207.

[38] M. Gronau, D. Pirjol, A. Soni, and J. Zupan, *Improved method for CKM constraints in charmless three-body B and B(s) decays*, Phys. Rev. **D75** (2007) 014002, arXiv:hep-ph/0608243.

[39] *Heavy Flavor Averaging Group*, http://www.slac.stanford.edu/xorg/hfag/.

[40] CDF Collaboration, M. J. Morello, *Charmless b−hadrons decays at CDF*, arXiv:0810.3258.

[41] A. Ali *et al.*, *Charmless non-leptonic B_s decays to PP, PV and VV final states in the pQCD approach*, Phys. Rev. **D76** (2007) 074018, arXiv:hep-ph/0703162.

[42] F. Su, Y.-L. Wu, C. Zhuang, and Y.-B. Yang, *Charmless $B_s \to PP, PV, VV$ Decays Based on the Six-Quark Effective Hamiltonian with Strong Phase Effects II*, Eur. Phys. J. **C72** (2012) 1914, arXiv:1107.0136.

[43] L. Evans, *LHC Machine*, Journal of Instrumentation **3** (2008) S08001.

[44] ATLAS Collaboration, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) S08003.

[45] ALICE collaboration, K. Aamodt *et al.*, *The ALICE Experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) S08004.

[46] CMS Collaboration, S. Chatrchyan *et al.*, *The CMS Experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) S08004.

[47] LHCb Collaboration, A. A. J. Alves *et al.*, *The LHCb Detector at the LHC*, Journal of Instrumentation **3** (2008) S08005.

[48] *lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/images/*, January 2013.

[49] *lbweb.cern.ch/groups/online/2011IntegratedLumiLHCbTime_NoPie.png*, September 2013.

[50] *lbweb.cern.ch/groups/online/2011AvgLumiTime.png*, September 2013.

[51] C. Lefevre, *The cern accelerator complex, CERN-DI-0606052*, 2006.

[52] LHCb Collaboration, *lhcb.web.cern.ch/lhcb/speakersbureau/html/*, January 2013.

[53] LHCb Collaboration, *Vertex locator technical design report, CERN-LHCC-2001-011*, .

[54] VELO working group, *lbtwiki.cern.ch/bin/view/VELO/VELOConferencePlots/*, January 2013.

[55] VELO working group, *lbtwiki.cern.ch/bin/view/VELO/MaterialInteractions*, January 2013.

[56] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, arXiv:physics/0503191.

[57] M. Needham, *Clone track identification using the kullback-liebler distance*, Tech. Rep. LHCb-2008-002. CERN-LHCb-2008-002. LPHE-2008-002, CERN, Geneva, Jan, 2008.

[58] V. Gligorov, C. Thomas and M. Williams, *The HLT inclusive b triggers, (CERN-LHCb-INT-2011-030)*, .

[59] R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055.

[60] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, Journal of Instrumentation **8** (2013) 2013P, arXiv:1210.6861.

[61] M. Pivk and F. R. Le Diberder, *Plots: A statistical tool to unfold data distributions*, Nuclear Instruments and Methods in Physics Research A **555** (2005) 356, physics/0.

[62] G. Punzi, *Sensitivity of Searches for New Signals and Its Optimization, arXiv:physics/0308063, (2003)*, .

[63] LHCb Collaboration, *Performance of the LHCb RICH detector at the LHC*, .

[64] T. Skwarnicki, *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*, .

[65] LHCb Collaboration, *Measurement of* $\sigma(\mathrm{pp} \to \mathrm{b\bar{b}X})$ *at* $\sqrt{s} = 7\,\mathrm{TeV}$ *in the forward region*, Phys. Lett. B **694** (2010) 209.

[66] L. Lyons, *Statistics for nuclear and particle physicists*, Cambridge University Press, Cambridge, 1986.

[67] L. Demortier and L. Lyons, *Everything you always wanted to know about pulls*, CDF/ANAL/PUBLIC/5776, 2002.

[68] K. S. Cranmer, *Kernel estimation in high-energy physics*, Comput. Phys. Commun. **136** (2001) 198, arXiv:hep-ex/0011057.

[69] *The Roofit toolkit for data modeling*, roofit.sourceforge.net/.

[70] *ROOT, an object oriented data analysis framework*, root.cern.ch/.

[71] C. J. Clopper and E. S. Pearson, *The use of confidence or fiducial limits illustrated in the case of the binomial*, Biometrika **26** (1934), no. 4 404.

[72] LHCb collaboration, R. Aaij *et al.*, *Search for* $B^{\pm} \to \phi\pi^{\pm}$ *decays and measurement of the charge asymmetry in* $B^{\pm} \to \phi K^{\pm}$, arXiv:1309.3742, submitted to Phys. Lett. B.

[73] J. Prisciandaro, F. Blanc, and O. Schneider, *Search for* $B^{\pm} \to \phi\pi^{\pm}$ *decays and measurement of the charge asymmetry in* $B^{\pm} \to \phi K^{\pm}$ *decays*, LHCb-ANA-2013-036.

[74] K. Cranmer, *Statistical Challenges for Searches for New Physics at the LHC*, in *Statistical Problems in Particle Physics, Astrophysics and Cosmology* (L. Lyons and M. Karagöz Ünel, eds.), p. 112, 2006. physics/0.

[75] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, Ann. Math. Statist. **9** (1938) 60.

[76] LHCb collaboration, *Updated average* f_s/f_d *b-hadron production fraction ratio for* 7 TeV *pp collisions*, LHCb-CONF-2013-011.

[77] E. O. Lebigot, *Uncertainties: a Python package for calculations with uncertainties, http://pythonhosted.org/uncertainties/*, 2013.

[78] R. H. Dalitz, *On the analysis of the* τ *meson data and the nature of the* τ *meson*, Phil. Mag. **44** (1953).

[79] *Search for CP violation in* $D^+ \to K^-K^+\pi^+$ *decays*, Phys. Rev. D **84** (2011).